

AJC:jcc
12/21/00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: Debra M. Eckert, David C. Chan, Vladimir N. Malashkevich, Peter A. Carr and Peter S. Kim

Continuation Application of

International Application No.: PCT/US99/17351

Filed: July 30, 1999

For: Inhibitors of HIV Membrane Fusion

Date: 12.21.00

EXPRESS MAIL LABEL NO. EL762340198US

REMARKS

Box PATENT APPLICATION

Assistant Commissioner for Patents

Washington, D.C. 20231

Sir:

The above-captioned application is a continuation of international application number PCT/US99/17351 filed on July 30, 1999 to which priority is claimed under 35 U.S.C. 120.

The related applications paragraph has been revised to include the parent application. No new matter has been added.

Respectfully submitted,
HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

By Anne J. Collins
Anne J. Collins
Registration No. 40,564
Telephone (781) 861-6240
Facsimile (781) 861-9540

Lexington, Massachusetts 02421-4799

Dated: December 21, 2000

AJC:jcc
December 21, 2000

-1-

Date: 12.21.00 Express Mail Label No. EL762340198US

Inventors: Debra M. Eckert, David C. Chan,
Vladimir N. Malashkevich, Peter A. Carr
and Peter S. Kim

Attorney's Docket No.: 0399.1192-008

INHIBITORS OF HIV MEMBRANE FUSION

RELATED APPLICATIONS

This application is a continuation of International Application No. PCT/US99/17351, entitled Inhibitors of HIV Membrane Infusion, by Debra M. Eckert, David C. Chan, Vladimir Malashkevich, Peter A. Carr and Peter S. Kim, filed July 30, 1999, which is related to U.S. Provisional Application 60/043,280, entitled Core Structure of gp41 from the HIV Envelope Glycoprotein, by David C. Chan, Deborah Fass, Min Lu, James M. Berger and Peter S. Kim, filed April 17, 1997 and U.S. Application 09/062,241, entitled Core Structure of gp41 from the HIV Envelope Glycoprotein, by David C. Chan, Deborah Fass, Min Lu, James M. Berger and Peter S. Kim, filed April 17, 1998. The present application claims the benefit of U.S. Provisional Application 60/094,676, entitled Inhibitors of HIV Membrane Fusion by David C. Chan, Debra M. Ehrgott and Peter S. Kim, filed July 30, 1998; U.S. Provisional Application 60/100,265, entitled Inhibitors of HIV Membrane Fusion, by David C. Chan, Debra M. Ehrgott and Peter S. Kim, filed September 14, 1998 and U.S. Provisional Application 60/101,058, entitled Inhibitors of HIV Membrane Fusion, by David C. Chan, Debra M. Ehrgott and Peter S. Kim, filed September 18, 1998; and U.S. Provisional Application 60/132,295, entitled Inhibitors of HIV Membrane Fusion, by Debra M. Ehrgott, David C. Chan, Vladimir Malashkevich and Peter S. Kim, filed May 3, 1999. The entire teachings of these referenced applications are incorporated herein by reference.

GOVERNMENT SUPPORT

The invention was supported, in whole or in part, by National Institutes of Health Grant Number P01 GM56552. The United States Government has certain rights in the invention.

5 BACKGROUND OF THE INVENTION

Structural studies of proteins from human immunodeficiency virus type 1 (HIV-1) have been essential in the development of anti-retroviral drugs. Structure-based drug development has been most intense for reverse transcriptase inhibitors and protease inhibitors, the two classes of HIV-1 drugs in clinical use. It would also be useful to be
10 able to carry out structure-based drug development against HIV entry.

SUMMARY OF THE INVENTION

As described herein, the cavities on the surface of the N-helix coiled-coil of HIV envelope protein gp41 subunit (e.g., HIV-1 envelope protein gp41-subunit) are targets for drugs or other agents which, by binding the coiled-coil surface, particularly the
15 cavities, inhibit HIV entry into cells. This is useful as the basis for identifying and designing drugs or agents which inhibit entry of HIV (e.g., HIV-1, HIV-2) into cells.

Results described herein show that the coiled-coil cavity (also referred to as the hydrophobic pocket) in the gp41 core is an attractive drug target and that molecules which bind the cavity interfere with (inhibit) HIV infectivity (HIV entry into cells).
20 Applicants have shown, for the first time, that conserved residues projecting into the hydrophobic pocket clearly play a major role in the ability of C34 to inhibit HIV-1 infection. The importance of cavity contacts (between the N-helix coiled-coil cavity and residues of the C peptide region of gp41) to gp41 function is clear. Conversely, the importance of preventing such cavity contacts in inhibiting gp41 function and, thus,

inhibiting HIV-1 entry into cells, is also clear. In addition, directing drugs against the hydrophobic pocket of the central-coiled coil of gp41 targets one of the most highly conserved regions of the HIV-1 envelope proteins, which means that drugs which target the coiled-coil surface, and particularly its hydrophobic pocket, will have broad activity
5 against diverse HIV isolates and that it will be difficult for drug-escape mutants to emerge.

A variety of methods, such as mirror-image phage display techniques (T. N. Schumacher, *et al.*, *Science*, 271:1854 (1996)), combinatorial chemistry (A. Borchardt, S. D. Liberles, S. R. Biggar, G.R. Crabtree, S.L. Schreiber, *Chem. Biol.*, 4:961 (1997);
10 J.C. Chabala, *Curr. Opin. Biotechnol.*, 6:632 (1995)), rational drug design and other drug screening and medicinal chemistry methods can be used to identify D-peptides, peptidomimetics and small molecules that bind the coiled-coil cavity with sufficient affinity to inhibit HIV-1 infection. The close correlation between N36/C34 stability and C34 potency, described herein, suggests that the effectiveness of such compounds will
15 depend critically on the strength of their cavity-contacts. As described herein, candidate compounds can be tested for their ability to interfere with formation of a stable complex between C34 and N36 or their ability to disrupt binding of the two (disrupt the complex), thereby providing rapid, quantitative screens to identify and evaluate potential inhibitors of HIV-1 entry.

20 Alternatively, screening can be carried out to identify molecules or compounds which interfere with or disrupt binding of the N-helix coiled-coil cavity and a peptide which binds the cavity, thus providing methods of identifying molecules which are “pocket specific” binding agents or drugs. Molecules and compounds described herein (also referred to as drugs or agents) are useful to inactivate gp41 and, thus, prevent or
25 reduce (inhibit) HIV-1 entry into cells. Without wishing to be bound by theory, it is reasonable to propose that these inhibitors bind to the pre-hairpin intermediate of gp41 and prevent its conversion to the trimeric hairpin structure of the gp41 core which corresponds to the fusion-active state of gp41. (Chan, D.C. and P.S. Kim, *Cell*, 93:681 (1998), See Figure 1). Thus, the present methods are useful to identify drugs or agents

which inhibit (totally or partially) formation of the fusion-active state of HIV-1 gp41 envelope protein. In the method, the ability of a candidate inhibitor (also referred to as a candidate drug), which can be any type of compound or molecule, such as a small molecule (e.g., a small organic molecule), a peptide (a D-peptide or an L-peptide), a peptidomimetic, a protein or an antibody, to bind the N-helix coiled-coil of gp41 and form a stable complex is assessed. Compounds or molecules which bind to the N-helix coiled-coil are further assessed for their ability to inhibit gp41 function (inhibit membrane fusion), such as through HIV-1 infection (viral entry) and syncytium assays, representative models of which are described and referenced herein. Those agents shown to inhibit gp41 function through such assays can be further assessed for their activity in additional *in vitro* assays and in appropriate animal models (e.g., Letvin, N.L., *Science*, 280, (5371): 1875 - 1880 (1998), Hirsch, V.M. and P.R. Johnson, *Virus Research*, 32 (2): 183-203 (1994); Reimann, K.A. *et al.*, *J. Virol.*, 70 (10): 6922-6928 (1996)). Any suitable approach can be used to assess binding of candidate inhibitors to the N-helix coiled-coil and, as a result of the work described herein, to the N-helix coiled coil cavity. In one embodiment, the ability of a candidate inhibitor to bind the synthetic peptide N36 (described in Lu, M. *et al.*, *J. Biomol. Struct. Dyn.* 15: 465 (1997), Chan, D.C. *et al.*, *Cell*, 89, 263 (1997) and U.S. Provisional Application 60/043,280, entitled Core Structure of gp41 From the HIV Envelope Glycoprotein, by David C. Chan, Deborah Fass, Min Lu, James M. Berger and Peter S. Kim, filed April 17, 1997) is assessed. The stability of the resulting complexes is assessed using methods described herein.

In a particular embodiment of the method of identifying compounds or molecules (drugs or agents) which bind the N-helix coiled-coil cavity, a soluble model that presents the gp41 coiled-coil cavity is used. The six helix bundle of HIV gp41 consists of an internal trimeric coiled-coil, composed of three identical N-peptides, surrounded by three C-peptides which fit into a conserved hydrophobic groove on the outside of the trimeric coiled-coil. The C-terminal end of the trimeric coiled-coil contains a large cavity into which bulky hydrophobic groups from the C-peptide pack.

This hydrophobic pocket is used as the target for anti-HIV drug discovery and/or design. Unfortunately, in the absence of the C-peptide, the N-peptide is aggregated and not 100% helical. Thus, simply using an N peptide from HIV-1 gp41, such as N36, N51 (Lu, M. *et al.*, *Nature Struct. Biology*, 1995) or DP-107 (Wild *et al.*, *PNAS* 89:10537-10541 (1992) is unlikely to provide an effective model for the N-helix coiled-coil.

As described herein, Applicants have succeeded in producing a soluble, non-aggregating trimeric peptide model of the hydrophobic pocket of HIV gp41 and, thus, for the first time, have provided a model that properly presents this hydrophobic pocket or cavity (in a manner or configuration which forms a similar structure to the corresponding residues in the HIV gp41 structure). (The terms “pocket” and “cavity” are used interchangeably.) As described, a peptide (also referred to as a fusion protein) which includes a soluble, trimeric coiled coil portion and a portion from the N-peptide region of HIV gp41 that includes the amino acid residues which form the pocket or cavity of the N-helix coiled-coil of HIVgp41 (the pocket-comprising residues of the N-peptide) has been produced and shown to be such a soluble model, useful to identify molecules or compounds which inhibit HIV gp41 function and, thus, HIV entry into cells. The trimeric version of the coiled-coil in the peptide (also referred to as a fusion protein) can be the coiled-coil region of a protein which is not a protein of HIV (a non HIV protein, such as GCN4-pIQI) or a protein of HIV origin (a protein derived from HIV or having the same or a similar amino acid sequence as an HIV protein). In a specific embodiment, the soluble, non-aggregating trimeric peptide model of the large cavity, referred to as IQN17, comprises a trimeric version of the coiled-coil region of GCN4, the yeast transcription activator, and a portion of the C-terminal end of the N peptide of gp41. IQN17 contains 29 residues of GCN4-pIQI (formerly referred to as GCN4-pIQ in U.S. Provisional Application 60/101,058) (Eckert, D.M. *et al.* *J. Mol. Biol.*, 284:859-865 (1998)), including three mutations for increased solubility, and 17 residues of HIV; there is a one residue overlap between the two proteins, making the total length of the fusion protein 45 residues. The sequence of GCN4-pIQI is: ac-RMKQIEDKIEEI LSKQYHIENEIAR IKKLIGER (SEQ ID NO:1). The HIV Sequence is: LLQLTVWG

IKQLQARIL (SEQ ID NO:20). The sequence of IQN17 is: ac-
 RMKQIEDKIEEIESKQKKIENEIARIKK LLQLTVWGIKQLQARIL-am (SEQ ID
 No:2). The HIV portion is underlined in SEQ ID No: 2; ac- represents an N-terminal
 acetyl group and -am represents a C-terminal amide. The sequence of the soluble,
 5 trimeric version of the coiled-coil region of GCN4 (referred to as a soluble, trimeric
 coiled coil of GCN4) in IQN17 is: RMKQIEDKIEEIESKQKKIENEIARIKK (SEQ ID
 No: 25). The superhelix parameters such as rise and pitch (Harbury, P.B. *et al.*, *Nature*
371:80-83 (1994); Harbury *et al.*, *PNAS* 92:8408-8412 (1995)) of the GCN4-pI_QI coiled
 coil are nearly identical to the HIV gp41 N-helix coiled coil. Therefore, the resulting
 10 fusion protein molecule (IQN17) is predicted to form a long trimeric coiled coil, which
 presents the N-peptide hydrophobic cavity at the C terminus. IQN17 is fully helical, as
 determined by circular dichroism, with a molar ellipticity at 222nm of -36,000 deg cm²
 dmol⁻¹. As determined by sedimentation equilibrium, IQN17 is close to a discrete
 trimeric species with a ratio of observed molecular weight to calculated molecular
 15 weight ranging from 3.00 to 3.16 times the monomer molecular weight at a
 concentration of 20 μM. As determined by X-ray crystallography, IQN17 presents the
 N-peptide hydrophobic pocket in a manner that is nearly identical to the pocket in the
 HIV gp41 N-helix coiled coil.

The IQN17 molecule (in the natural L-handedness or enantiomeric D-
 20 handedness) can be used in screens, including high-throughput drug screens, to identify
 molecules that bind to the coiled-coil pocket. The IQN17 molecule, in the D-
 handedness, has been used as a target in mirror image phage display (Schumacher *et al.*,
Science, 271: 1854, 1996) to identify small molecules (D-peptides) which bind to the
 hydrophobic pocket of gp41 (in the natural L-handedness) and inhibit HIV-membrane
 25 fusion. The desired target (the N-helix of HIV gp41 which includes the hydrophobic
 pocket) is the exact mirror image of the naturally-occurring target. It is used to screen a
 library or collection of compounds or molecules which are to be assessed for their
 ability to bind the mirror image of the naturally-occurring coiled-coil pocket. The mirror
 image of a compound or molecule found to bind the mirror image of the naturally-

occurring gp41 pocket, will bind the gp41 pocket in the natural handedness. The library or collection screened can be of any type, such as a phage display library, peptide library, DNA library, RNA library, combinatorial library, collection of chemical agents or drugs, cell lysate, cell culture medium or supernatant containing products produced by cells. In the case of a phage display library, the D-target is used to screen phage coat proteins. Specific phage clones that bind to the target are identified and the mirror images of the expressed proteins are chemically synthesized with D-amino acids. By using IQN17 in mirror-image phage display, D-peptides that bind to the gp41 hydrophobic pocket have been identified. Further assessment has been carried out, as described, to demonstrate the ability of D-peptides to inhibit HIV gp41 function. D-peptides which bind the gp41 hydrophobic pocket and inhibit HIV infectivity have been identified. D-peptides which bind the hydrophobic pocket also will serve as lead molecules for drug development and/or reagents for drug discovery (where the drugs bind to the coiled-coil pocket and inhibit HIV infectivity). The IQN17 molecule, in the natural L-handedness, can be used in screens, including high-throughput screens, to identify molecules that bind to the coiled-coil pocket. IQN17 can be used to screen a collection or library of compounds or molecules which are to be assessed for their ability to bind the hydrophobic pocket. The library or collection screened can be of any type, such as a phage display library, RNA library, DNA library, peptide library, combinatorial library, collection of chemical agents or drugs, cell lysate, cell culture medium or supernatant containing products produced by cells. Compounds or molecules which bind the hydrophobic pocket also will serve as lead molecules for drug development and/or reagents for drug discovery.

Fusion proteins which are variants of IQN17 can be produced and used to screen for drugs which bind the gp41 N-helix coiled-coil pocket. Any of a wide variety of variations can be made in the GCN4-pI_QI component of IQN17 and used in the method, provided that these changes do not alter the trimeric state of the coiled-coil. For example, the amino acid composition of the GCN4 component can be changed by the addition, substitution, modification and/or deletion of one or more amino acid residues,

provided that the trimeric state of the coiled-coil is maintained. For example, the Asp residue in IQN17 (at a “f-position” of the coiled coil) can be replaced by any of the naturally-occurring amino acids. (O’Neil and DeGrado, *Science* 250:646 (1990)). Alternatively, this component of the fusion protein can be a trimeric version of the coiled-coil region of another protein, such as that from Moloney Murine Leukemia Virus (Fass, D. *et al. Nature Struct. Biology*, 3:465 (1996)), GCN4-pII (Harbury *et al., Nature*, 317:80, 1994) or the ABC heterotrimer (Nautiyal and Alber, *Protein Science* 8:84 (1999)).

Changes can also be made in the amino acid composition of the fusion protein component which is the C-terminal portion of the HIV gp41 N peptide to produce IQN17 variants. The C-terminal portion can be changed by the addition, substitution, modification and/or deletion of one or more amino acid residues. The amino acid composition of either or both components of the fusion protein can be altered, and there is no limit to the number or types of amino acid residue changes possible, provided that the trimeric state of the coiled-coil and the hydrophobic pocket of the N peptide of HIV gp41 are maintained. IQN17, IQN17 variants or any soluble model of the large cavity can be used to screen for drugs which bind the N-helix coiled-coil, especially the pocket, or for lead drug candidates or candidates for use in vaccine preparations, to be further screened using methods known to those of skill in the art, such as in a high throughput format.

Results described herein are useful to screen for inhibitors of HIV gp41 which are variants of C34 as described below. Once a variant of C34, such as a C34 variant which stably binds N36, has been identified, it can be used and further assessed as obtained or it can be modified (e.g., by altering, adding, deleting or substituting at least one amino acid residue or adding a non-amino acid substituent), if desired or needed (e.g., to enhance stability, solubility, bioavailability). Alternatively, a C34 variant can be assessed to determine if a shorter component (region of fewer amino acid residues) also is active as an inhibitor. As discussed herein, the three C34 residues Trp⁶²⁸, Trp⁶³¹ and Ile⁶³⁵ that pack into the deep, conserved pocket in the N36 trimer are critical for

inhibitory activity. The observation that C34 variants that have a higher affinity for the N36 coiled-coil have more potent inhibitory activity against HIV infection forms the basis for screens to identify and evaluate potential inhibitors. For example, using the "split-synthesis" technique (Chen, C.L., *et al. Methods Enzymol.* 267:211-219 (1996);

5 Lam, K.S. *et al.*, *Nature*, 354: 82-84, (1991)) of combinatorial peptide chemistry, a library of C34 variants is synthesized in which the three critical hydrophobic residues are randomly replaced by chemical substitutions of varying hydrophobic character. This synthesis technique results in the generation of a vast library of beads, each containing many copies of a single variant C34 peptide (i.e., a "one-bead, one-compound" type of

10 library). To identify C34 variants which stably bind the N-helix coiled-coil, a labeled version of N36 (or a modified N-peptide) is mixed with the peptide beads under conditions (e.g., elevated temperature) that restrict binding to only those C34 variants with the highest affinity. Binding is measured by detection of the label on the N-helix peptide, using known methods. Simple modifications of the split-synthesis technique

15 allow ready identification of the selected peptide sequence by mass spectroscopy (Youngquist, R.S. *et al.*, *J. Amer. Chem. Soc.* 117, 3900-3906 (1995)). The C34 variants selected, particularly those with the highest binding affinities for N36, are tested in syncytium and infection assays for gp41 inhibitory activity. Truncated versions of these C34 variants, containing only the cavity-binding region, can also be tested for

20 inhibitory activity. Alternatively, a library of other peptides to be assessed can be synthesized to generate a library of beads, each containing (having bound thereto) a peptide to be assessed. This library is analyzed as described above for the C34 variants and resulting hits (members with appropriate binding affinities for N36) are further analyzed for gp41 inhibitory activity. As a second example, the N36 peptide or the

25 soluble variants described earlier, such as IQN17, GCN4-N-helix peptide can be used as a target for phage display or mirror-image phage display techniques to identify peptides that bind to the cavity.

IQN17 can also be used to raise antibodies (monoclonal and/or polyclonal) that bind to the coiled-coil cavity. IQN17 can further be used, either alone or in combination

with other materials, in a vaccine, which will elicit the production of antibodies that bind to the coiled-coil in the individual to whom it is administered (the vaccinee), and thereby offer protection against infection and/or disease.

Peptides, both D-peptides and L-peptides, which fit into a deep hydrophobic pocket in the trimeric N-helix coiled-coil of HIV-1 envelope glycoprotein gp41 are also the subject of this invention. The D-peptides are the first molecules that have been shown to bind exclusively to the gp41 hydrophobic pocket. The observation that these D-peptides inhibit gp41-mediated membrane fusion processes (syncytia formation and viral infection) provides the first direct demonstration that HIV-1 infection can be inhibited by molecules that bind specifically to pocket. The validation of the gp41 hydrophobic pocket as a drug target sets the stage for the development of a new class of orally bioavailable anti-HIV drugs, that work by inhibiting viral entry into cells. Such drugs would be a useful addition to the current regimen used to treat HIV-1 infection with combination therapies. D-peptides, such as the D-peptides described herein, portions, modification and variants thereof and larger molecules (e.g., polypeptides) which comprise all or a portion of a D-peptide described herein, are useful to inhibit HIV membrane fusion and, thus, HIV entry into cells. D-peptides, corresponding to the D-amino acid version of phage sequences identified as described herein, are inhibitors of HIV-1 infection and syncytia formation. The C-terminal residues in these D-peptide inhibitors have the sequence pattern: CXXXXXEWXWLCAA-am. (In the phage-display library, the positions corresponding to the C residues were encoded as either C or S, the positions corresponding to the AA residues were encoded as such and the other 10 positions (indicated by X) were randomly encoded. The -am represents a C-terminal amide, added as part of the peptide synthesis procedure.) The N-terminal residues in the D-peptide inhibitors are, for example, ac-GA, ac-KKGA, or ac-KKKKGA. The ac- represents an N-terminal acetyl group added as part of the peptide synthesis procedure. The C-terminal amide and the N-terminal acetyl group are optional components of D-peptides of this invention. Other N-terminal residues can be included, in place of or in

addition to those in the previous sentence, as desired (e.g., to increase solubility). For example, D-peptides of the following sequences are also the subject of this invention:

ac-XXCXXXXXEWXWLCXX-am (SEQ ID NO: 28);
 ac-KKXXCXXXXXEWXWLCXX-am (SEQ ID NO: 29);
 5 ac-KKKKXXCXXXXXEWXWLCXX-am (SEQ ID NO: 30);
 ac-XXCXXXXXEWXWLCXXX-am (SEQ ID NO: 31);
 ac-KKXXCXXXXXEWXWLCXXX-am (SEQ ID NO: 32); and
 ac-KKKKXXCXXXXXEWXWLCXXX-am (SEQ ID NO: 33).

The amino acid residues are represented by the single letter convention and X
 10 represents any amino acid residue (naturally occurring or non-naturally occurring) or other moiety, such as a modified amino acid residue.

Further, the ten amino acid residue "core" (the 10-mer which is flanked at each
 end by a cysteine residue) of the 12 amino acid residue peptide, as well as portions,
 modifications and variants of the 10-mers are also useful to inhibit membrane fusion
 15 and entry of HIV into cells. Variants, portions and modifications of these peptides are also useful as inhibitors. As described further herein, D-peptides which comprise a consensus sequence (e.g., WXWL (SEQ ID NO: 23), EWXWL (SEQ ID NO: 24), CXXXXXEWXWLC (SEQ ID NO: 12) or a portion thereof) have been shown to bind
 the N-helix coiled-coil and are useful to inhibit membrane fusion and entry of HIV into
 20 cells. The enantiomeric peptides (D-peptides) do not serve as efficient substrates for enzymes, such as proteases and, therefore, are more resistant to proteolytic degradation than are L-peptides; they are also less immunogenic than are L-peptides.

Specific embodiments of D-peptides of the present invention are:

- (a) CDLKAKEFWWLC (SEQ ID NO: 3);
- 25 (b) CEARHREWAWLC (SEQ ID NO: 4);
- (c) CELLGWEWAWLC (SEQ ID NO: 5);
- (d) CLLRAPEWGWLC (SEQ ID NO: 6);
- (e) CSRSQPEWEWLC (SEQ ID NO: 7);

- (f) CGLGQEEFWWLC (SEQ ID NO: 8);
- (g) CMRGEWEWSWLC (SEQ ID NO: 9);
- (h) CPPLNKEWAWLC (SEQ ID NO: 10);
- (i) CVLKAKEFWWLC (SEQ ID NO: 11);
- 5 (j) KKGACGLGQEEFWWLC (SEQ ID NO: 15);
- (k) KKGACCELLGWEWAWLC (SEQ ID NO: 16);
- (l) KKKKGACCELLGWEWAWLC (SEQ ID NO: 17);
- (m) KKGACMRGEWEWSWLC (SEQ ID NO: 18);
- (n) KKGACPPLNKEWAWLC (SEQ ID NO: 19);
- 10 (o) a D-peptide comprising WXWL (SEQ ID NO: 23);
- (p) a D-peptide comprising EWXWL (SEQ ID NO: 24);
- (q) a D-peptide comprising CXXXXXEWXWL (SEQ ID NO: 12)
- (r) ac-GACEARHREWAWLCAA-am (SEQ ID NO: 34);
- (r) ac-KKGACEARHREWAWLCAA-am (SEQ ID NO: 38);
- 15 (t) ac-KKKKGACEARHREWAWLCAA-am (SEQ ID NO: 43);
- (u) ac-GACGLGQEEFWWLCAA-am (SEQ ID NO: 44);
- (v) ac-KKGACGLGQEEFWWLCAA-am (SEQ ID NO: 15);
- (w) ac-KKKKGACGLGQEEFWWLCAA-am (SEQ ID NO: 45)
- (x) ac-GACDLKAKEFWWLCAA-am (SEQ ID NO: 35);
- 20 (y) ac-KKGACDLKAKEFWWLCAA-am (SEQ ID NO: 39);
- (z) ac-KKKKGACDLKAKEFWWLCAA-am (SEQ ID NO: 46);
- (a') ac-GACCELLGWEWAWLCC-am (SEQ ID NO: 47);
- (b') ac-KKGACCELLGWEWAWLCAA-am (SEQ ID NO: 16);
- (c') ac-KKKKGACCELLGWEWAWLCAA-am (SEQ ID NO: 17);
- 25 (d') ac-GACSRSQPEWEWLCAA-am (SEQ ID NO: 36);
- (e') ac-KKGACSRSQPEWEWLCAA-am (SEQ ID NO: 40);
- (f') ac-KKKKGACSRSQPEWEWLCAA-am (SEQ ID NO: 48);
- (g') ac-GACLLRAPEWGWLCAA-am (SEQ ID NO: 37);
- (h') ac-KKGACLLRAPEWGWLCAA-am (SEQ ID NO: 41);

- (i') ac-KKKKGACLLRAPEWGWLCAA-am (SEQ ID NO: 49);
- (j') ac-GACMRGEWEWSWLCAA-am (SEQ ID NO: 50);
- (k') ac-KKGACMRGEWEWSWLCAA-am (SEQ ID NO: 18);
- (l') ac-KKKKGACMRGEWEWSWLCAA-am (SEQ ID NO: 51);
- 5 (m') ac-GACPPLNKEWAWLCAA-am (SEQ ID NO: 52);
- (n') ac-KKGACPPLNKEWAWLCAA-am (SEQ ID NO: 19);
- (o') ac-KKKKGACPPLNKEWAWLCAA-am (SEQ ID NO: 53);
- (p') ac-GACXXXXXEWXWLCAA-am (SEQ ID NO: 54);
- (q') ac-KKGACXXXXXEWXWLCAA-am (SEQ ID NO: 55);
- 10 (r') ac-KKKKGACXXXXXEWXWLCAA-am (SEQ ID NO: 56);
- (s') ac-XXCXXXXXEWXWLCXX-am (SEQ ID NO: 57);
- (t') ac-KKXXCXXXXXEWXWLCXX-am (SEQ ID NO: 58);
- (u') ac-KKKKXXCXXXXXEWXWLCXX-am (SEQ ID NO: 59);
- (v') ac-XXCXXXXXEWXWLCXXX-am (SEQ ID NO: 60);
- 15 (w') ac-KKXXCXXXXXEWXWLCXXX-am (SEQ ID NO: 61);
- (x') ac-KKKKXXCXXXXXEWXWLCXXX-am (SEQ ID NO: 62); and
- (y') a variant of a sequence of (a) through (x'), wherein the variant binds the N-helix coiled-coil cavity of HIV gp41, wherein ac- at the C-terminus and -am at the N-terminus are optional.

20 D-peptides described herein, which are ligands shown to bind the N-helix pocket, are also useful in drug screens to identify compounds or molecules (e.g., from chemical libraries, recombinantly produced products, naturally-occurring substances, culture media or supernatants) which bind the N-helix pocket and thus, are also inhibitors of HIV. For example, a competitive assay can be carried out by combining a

25 D-peptide which binds the N-helix cavity (e.g., a D-peptide described herein); IQN17 (e.g., in the natural L-handedness), or another fusion protein which is a soluble model that presents the N-helix cavity; and a candidate inhibitor (a compound or molecule to be assessed for its ability to bind the N-helix cavity). For example, D10pep5 or

D10pep1, IQN17, and a candidate inhibitor (candidate drug) can be combined using buffer conditions and peptide concentrations appropriate for binding of D10pep5 or D10pep1 to IQN17. The extent to which binding of the D-peptide occurs is determined and compared to the extent to which binding occurs under the same conditions, but in the absence of a compound or molecule (referred to as a candidate drug or candidate inhibitor) to be assessed for its ability to bind the N-helix coiled-coil cavity of HIV gp41 envelope protein (in a control). If binding of D10pep5 or D10pep1 occurs to a lesser extent in the presence of the candidate inhibitor (test sample) than in its absence (control sample), the candidate inhibitor is a ligand which binds the N-helix coiled-coil cavity and, thus, is an inhibitor. Inhibitors identified in this manner can be further assessed for their activity in viral infectivity assays and syncytia formation assays, such as those described herein. Those inhibitors which show activity in such assays can be further assessed in an appropriate animal model or in humans.

Any method by which binding of the D-peptide, known to bind the N-helix cavity, can be detected can be used to assess whether the candidate inhibitor interferes with binding. For example, the D-peptide can be detectably labeled and the extent to which the label appears on the N-helix cavity (as a result of binding of the D-peptide) detected, in the presence and in the absence of the candidate inhibitor. If less label appears on the N-helix cavity of IQN17 (or other appropriate fusion protein) in the presence of the candidate inhibitor (in the test sample) than in its absence (in the control sample), then the candidate inhibitor is a ligand which binds the N-helix cavity (and interferes with binding of the D-peptide). Alternatively, the D-peptide (e.g., D10pep5 or D10pep1) and IQN17 can be labeled with a fluorophore (e.g., with EDANS; 5-(2'-aminoethyl)aminonaphthalene-1-sulfonic acid) with an appropriate quencher that quenches the fluorescent signal of the fluorophore when it is in close proximity (e.g., DABCYL; 4-(4'-dimethylaminophenylazo)benzoic acid). If the candidate inhibitor binds the N-helix cavity of IQN17, fluorescence is observed, since, as a result of binding of the candidate inhibitor, the D-peptide is not brought into sufficiently close proximity to the quencher to permit it to quench the reporter signal. Alternatively, the

fluorescent reporter molecule can be on the IQN17 and an appropriate quencher on the D-peptide. In either case, the position of the reporter or quencher on IQN17 must be such that when the D-peptide binds the N-helix cavity, the reporter and quencher moieties are in sufficiently close proximity to each other that quenching occurs (Tyagi, S., *et al.*, *Nature Biotechnology* 16:49 (1998)).

Also the subject of this invention are drugs (compounds and molecules) which bind the N-helix coiled-coil pocket of HIV gp41 and inhibit (partially or totally) HIV entry into cells. In one embodiment, these drugs can be identified as described herein or by other methods. Drugs which bind the N-helix coiled-coil pocket of HIV gp41 are useful as therapeutic agents (to prevent HIV entry into cells or reduce the extent to which it occurs), as research tools (e.g., to study the mechanism of HIV gp41 function) and to assess the rate of viral clearance by an individual (e.g., in an animal model or an infected human).

Also the subject of this invention are compositions, useful in methods of interfering with entry of HIV into a mucosal cell; these compositions comprise an appropriate carrier or base and at least one component selected from the group consisting of:

- (a) C34 peptide;
- (b) DP178;
- (c) T649;
- (d) T1249;
- (e) a derivative of (a) - (d);
- (f) a D-peptide which binds to the hydrophobic pocket of HIV gp41;
- (g) a derivative of (f);
- (h) a combination of two or more of (a)-(g); and
- (i) a molecule that inhibits HIV infectivity by binding to the N-helix coiled coil.

The compositions can comprise one such component or two or more components.

A further subject of this invention are compositions (e.g., proteins or proteinaceous materials) that can be used to elicit an immune response (e.g., antibody production) that will protect (partially or totally) against HIV infection and/or disease. Such compositions are useful as protective agents (e.g., vaccines) and to obtain

5 antibodies (monoclonal and/or polyclonal) that are useful as research tools, diagnostic tools, drug screening reagents, and to assess viral dynamics (rates of production and clearance of virus) in animal models or infected humans.

Also the subject of this invention is a list of atomic coordinates for the X-ray crystal structure of the complex between IQN17 and D10pep1. Also the subject of this

10 invention is a list of coordinates for the X-ray crystal structure of IQN17. These coordinates can be used (e.g., as an electronic file for computer graphics programs) to create a model of the complex which indicates how D10pep1 binds to the N-helix coiled-coil cavity and models of the N-helix coiled-coil cavity. Such models can be used, in methods known to those of skill in the art such as in computer graphics

15 modeling, to build new models to evaluate the likelihood of binding to the N-helix coiled-coil cavity by other peptides, peptidomimetics, small molecules, drugs or other compounds. Such models can also be used to build new models for the structures of molecules (peptides, peptidomimetics, small organic molecules, drugs or other compounds) that bind the N-helix coiled-coil cavity (e.g., H. Kubinyi, *Curr. Op. Drug*

20 *Discov. Develop.*, 1:16 (1998); P.L. Wood, *ibid*, 1:34 (1998); J.R. Morphy, *ibid*, 1:59 (1998)). These models and the corresponding lists of atomic coordinates can be used to identify, evaluate, discover and design more effective and/or new D-peptides, L-peptides, peptidomimetics, other small molecules or drugs that inhibit HIV infectivity, using methods known to those of skill in the art. A further subject of this invention is a

25 method of producing or identifying a drug which fits (packs into, binds) the N-helix coiled-coil pocket of HIV gp41 through the use of atomic coordinates of a crystal, such as a crystal of a soluble, trimeric peptide model of the HIV gp41 hydrophobic pocket described herein (e.g., IQN17 or a variant thereof), a crystal of such a model in complex with a D-peptide (e.g., IQN17 or a variant thereof in complex with a D-peptide

described herein, such as D10pep1) or a crystal of the N-peptide region of HIV gp41 comprising the amino acid residues which comprise the pocket of the N-helix coiled-coil of HIV gp41. The method comprises obtaining a crystal of the soluble model, such as the empty soluble model (not in complex with a D-peptide), obtaining the atomic

5 coordinates of the crystal (e.g., of the crystal of the empty soluble model, such as IQN17); using the atomic coordinates obtained to define the N-helix coiled-coil pocket of HIV gp41; identifying a molecule or compound which fits the N-helix coiled-coil pocket and obtaining the molecule or compound; contacting the molecule or compound with the N-helix coiled-coil pocket (e.g., by contacting it with a polypeptide which

10 comprises the pocket (e.g., IQN17 or a variant thereof or the N-peptide) to assess (determine) the ability of the molecule or compound to fit the pocket of HIV gp41, wherein in the molecule or compound fits the pocket, it is a drug which fits the N-helix coiled-coil pocket, whereby a drug which fits the pocket is produced. The atomic coordinates of the crystal can be obtained by X-ray diffraction studies or from a

15 computer file or Protein Data Base (PDB), such as the PDB presented herein for IQN17 (Figures 11A-11V).

Similarly, the method can be carried out using a crystal of a soluble trimeric model in complex with a D-peptide (e.g., a D-peptide described herein, such as D10pep1) or a crystal of the N-peptide region of HIV gp41 which comprises the pocket

20 of the N-helix coiled coil.

Drugs produced in this manner can be further assessed to conform their ability to fit into the pocket (e.g., by NMR) and can be assessed for their ability to inhibit HIV entry into cells (e.g., by a syncytia assay or infectivity assay).

The teachings and entire contents of all documents cited herein are expressly

25 incorporated by reference into this application.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic of HIV-1 gp41 showing the N36 (SGIVQQQNNLLRAIEQQHLLQLTVWGIKQLQARIL) (SEQ ID NO:13) and C34

(WMEWDREINNYTSLIHSLIEESQNQQEKNEQELL) (SEQ ID NO: 14) peptides located within two regions containing 4,3 hydrophobic heptad repeats (labeled heptad repeat 1 and heptad repeat 2, also referred to as N-peptide region and C-peptide region, respectively). The underlined residues in C34 were mutated in this study. Three of
 5 these residues (W, W and I) project into the N36 cavity, whereas two of these residues (M and R) do not. FP, fusion peptide; S-S, disulfide bond; TM, transmembrane region; INTRA, intraviral region.

Figure 2 is a graph showing the correlation of C34 inhibitory potency with N36/C34 stability. C34 peptide variants containing substitutions at position Trp⁶³¹ were
 10 tested for inhibition of viral entry (filled circles) and cell-cell fusion (open circles). IC₅₀ values are plotted on a logarithmic scale against the T_m (melting temperature) of the corresponding N36/C34 complex. The identities and chemical structures of the substitutions are drawn under the corresponding data points. In order of increasing hydrophobic bulk, the substitutions were: glycine (Gly), alanine (Ala), L- α -
 15 aminobutyric acid (Abu), valine (Val), leucine (Leu), phenylalanine (Phe), the wildtype residue tryptophan (Trp), and L- β -(1-naphthyl) alanine (Nal). Error bars indicate the standard error from triplicate experiments.

Figure 3 shows the amino acid sequences of D-peptides (SEQ ID NOS: 34, 38, 15, 35, 16, 17, 36, 40, 41, 18 and 19) and the consensus sequence (SEQ ID NO.: 12).
 20 As represented, each peptide is flanked by GA on the N-terminus and AA on the C-terminus, and comprises a blocking group at the N-terminus: (Acetyl-GA-C-10mer-C-AA-CONH₂; this can also be represented as ac-GA-C-10mer-C-AA-am). The single letter conventions which are used to represent amino acid residues are as follows:
 G=glycine; A=alanine; C=cysteine; D=aspartic acid; L=leucine; K=lysine; E=glutamic
 25 acid; W=tryptophan; F=phenylalanine; R=arginine; H=histidine; S=serine; and
 Q=glutamine.

Figure 4 is a schematic representation of mirror-image phage display with the D-IQN17 target, in which: (1) rounds of phage selection are carried out to identify binders to D-IQN17; (2) individual clones are sequenced; (3) binding specificity is assessed by

determining whether the phage bind to the gp41 region of D-IQN17; (4) D-peptides of those phage sequences which bind are produced; and (5) the anti-HIV activity of the D-peptides is assayed.

Figures 5A and 5B show the crystal structure of IQN17 bound to D10pep1.

- 5 IQN17 is shown to be a continuous three-stranded coil, and binding of the conserved amino acid residues of D10pep1 is shown to be to the hydrophobic pocket of IQN17, formed by the 17 residues derived from HIV gp41. Figure 5A shows IQN17, consisting of GCN4-pIQI residues fused to HIV-1 gp41 residues and the binding of D10pep1 to the hydrophobic pocket of IQN17 (area within box). The D-peptide which binds to the
- 10 pocket is represented by the branched extensions (i.e., stick representation). Figure 5B is an enlargement of the area within the box and shows the conserved residues that pack into the pocket (Trp, Trp Leu) as well as a glutamic acid (Glu).

Figures 6A and 6B show results of syncytia assays, using the D-peptides described herein. Figure 6A is a graphic representation of results of syncytia assays.

- 15 Figure 6B represents IC_{50} data for D-peptides, with results from one or more experiments.

- Figures 7A-7N are the PDB file which lists the atomic coordinates for the crystal structure of D10pep1 bound to IQN17, in which residues 0-28 of the A chain are derived from the GCN4-pIQI sequence (with three mutations), residues 29-45 of the A chain are derived from the HIVgp41 sequence, residues 0-16 of the D chain represent the D-peptide, ordered water molecules are represented as W, and a bound chloride ion as chain I. Residue 0 represents the acetyl group. The PDB file represents a monomer; the trimer is formed by crystallographic symmetry.
- 20

- Figures 8A and 8B show results of assessment of inhibition of HIV-1 membrane fusion by a D-peptide. Figure 8A shows results of syncytia assay with no D-peptide. Figure 8B shows results of syncytia assay with D-peptide.
- 25

Figures 9A-9C show results of 1H NMR experiments characterizing the aromatic residues of IQN17/D-peptide complexes. Figure 9A shows 1D-NMR spectra of D10pep1a (top), IQN17 (middle) and a 1:1 complex of D10pep1a and IQN17 (bottom).

The x-axis is the same as for (C) below. Upfield peaks assigned to the four scalar-coupled aromatic ring protons of Trp-571 are indicated. The unmarked upfield peak of the bottom trace corresponds to an unassigned H α resonance. Figure 9B shows 1D spectra of 1:1 complexes between IQN17 and each D-peptide (as labeled). The same
 5 four protons are indicated in some spectra. Figure 9C shows a 2D-NMR TOCSY spectrum of IQN17/D10pep1a complex. Cross-peaks linking these four tryptophan protons are indicated, along with specific assignments. The TOCSY mixing time was 42ms.

Figure 10 shows the conformation of the D10pep1 peptide as in the complex
 10 with IQN17, as determined by X-ray crystallography.

Figures 11A and 11V are the PDB file which lists the atomic coordinates for the crystal structure of IQN17, in which residues 0-28 of the A, B and C chains of the IQN17 trimer are derived from GCN4-pI_QI sequence (with three mutations), residues 29-45 of the chains A, B, and C are derived from HIV gp41, ordered water molecules
 15 are represented as W, and a bound chloride ion as chain I. The PDB file represents a whole trimer in the crystallographic asymmetric unit.

DETAILED DESCRIPTION OF THE INVENTION

The gp41 subunit of the HIV-1 envelope protein mediates fusion of viral and cellular membranes. The crystal structure of the gp41 ectodomain core is a six-helix
 20 bundle composed of three helical hairpins, each consisting of an N-helix paired with an antiparallel C-helix (D. C. Chan, D. Fass, J. M. Berger, P. S. Kim, *Cell*, 89:263 (1997), W. Weissenhorn, A. Dessen, S.C. Harrison, J. J. Skehel, D. C. Wiley, *Nature*, 387:426 (1997); K. Tan, J. Liu, J. Wang, S. Shen, M. Lu, *Proc. Natl. Acad. Sci. USA*, 94:12303 (1997). Three N-helices form an interior, trimeric coiled-coil, and three C-helices wrap
 25 around the outside of this N-helix coiled-coil along conserved, hydrophobic grooves. This structure likely corresponds to the core of the fusion-active state of gp41 (discussed in D. C. Chan, D. Fass, J. M. Berger, P. S. Kim, *Cell*, 89:263 (1997), and D.C. Chan and Peter S. Kim, *Cell*, 93:681 (1998)) and shows similarity to the proposed fusogenic

- structures of envelope fusion proteins from influenza (P. A. Bullough, F. M. Hughson, J. J. Skehel, D. C. Wiley, *Nature*, 371:37 (1994)), Moloney Murine Leukemia Virus (D. Fass, S. C. Harrison, P. S. Kim, *Nat. Struct. Biol.*, 3:465 (1996)), and simian immunodeficiency virus (SIV). (V.N. Malashkevich, D.C. Chan, C.T. Chutkowski, P.S. Kim, *Proc. Natl. Acad. Sci. USA*, 95:9134 (1998), M. Caffrey *et al.*, *EMBO J.*, 17:4572 (1998)), and Ebola virus (W. Weissenhorn *et al.*, *Mol. Cell* 2:605 (1998), V. N. Malashkevich *et al.*, *Proc. Natl. Acad. Sci. USA*, 96:2662 (1999).)
- Synthetic C-peptides (peptides corresponding to the C-helix), such as DP178 and C34, are potent inhibitors of HIV-1 membrane fusion and are effective against both laboratory-adapted strains and primary isolates (V. N. Malashkevich, D. C. Chan, C. T. Chutkowski, P. S. Kim, *Proc. Natl. Acad. Sci. USA*, 95:9134 (1998), DP178 corresponds to residues 638-673 of HIV-1 gp41 and is acetylated at the amino terminus and amidated at the carboxy terminus (C. T. Wild, D. C. Shugars, T. K. Greenwell, C. B. McDanal, T. J. Matthews, *Proc. Natl. Acad. Sci. USA*, 91:9770 (1994), S. Jiang, K. Lin, N. Strick, A.R. Neurath, *Nature*, 365:113 (1993)). A Phase I clinical trial with the C-peptide DP178 (also called T-20) indicates that it has antiviral activity *in vivo*, resulting in reduced viral loads (M. Saag, *et al.*, abstract #771 presented at the Infectious Disease Society of America 35th Annual Meeting, San Francisco, CA, 16 September 1997; Kilby, J.M. *et al.* *Nature Med.* 4:1302-1307 (1998)). Based on the structural features of the gp41 core, these peptides are thought to act through a dominant-negative mechanism, in which exogenous C-peptides bind to the central coiled-coil of gp41 and lead to its inactivation (D.C. Chan and P.S. Kim, *Cell*, 93:681 (1998); R.A. Furuta *et al.*, *Nat. Struct. Biol.*, 5:276 (1998); D. C. Chan, D. Fass, J. M. Berger, P. S. Kim, *Cell*, 89:263 (1997), W. Weissenhorn, A. Dessen, S.C. Harrison, J. J. Skehel, D. C. Wiley, *Nature*, 387:426 (1997); K. Tan, J. Liu, J. Wang, S. Shen, M. Lu, *Proc. Natl. Acad. Sci. USA*, 94:12303 (1997), M. Lu, S. C. Blacklow, P. S. Kim, *Nat. Struct. Biol.*, 2:1075(1995) and C. H. Chen, T. J. Matthews, C. B. McDanal, D. P. Bolognesi, M. L. Greenberg, *J. Virol.*, 69:3771 (1995)). These peptides likely act on a pre-hairpin intermediate of gp41 that forms when the native gp41 structure (i.e., the nonfusogenic

conformation present on free virions) is perturbed by gp120/CD4/coreceptor interactions. This pre-hairpin intermediate is proposed to have an exposed N-coiled-coil, thereby allowing C-peptides to bind and inactivate gp41 prior to the formation of the fusion-active hairpin structure (D. C. Chan, P. S. Kim, *Cell*, 93:681 (1998)). This model is further supported by immunoprecipitation experiments indicating that the C-peptide DP178 binds to gp41 (R. A. Furuta, C. T. Wild, Y. Weng, C. D. Weiss, *Nat. Struct. Biol.*, 5:276 (1998)). In addition, viruses escaping DP178 inhibition show mutations in the central coiled-coil region of gp41 (L. T. Rimsky, D. C. Shugars, T. J. Matthews, *J. Virol.*, 72:986 (1998)).

Recent crystallographic studies of gp41 facilitate the development of small-molecule peptidomimetic drugs which, in contrast to C-peptides, have the potential to be orally administered. Within each coiled-coil interface is a deep cavity, formed by a cluster of residues in the N-helix coiled-coil, that is an attractive target for the development of antiviral compounds. Three residues from the C-helix (Trp⁶²⁸, Trp⁶³¹, and Ile⁶³⁵) insert into this cavity and make extensive hydrophobic contacts. Mutational analysis indicates that two of the N-helix residues (Leu⁵⁶⁸ and Trp⁵⁷¹) comprising this cavity are critical for membrane fusion activity (J. Cao, *et al.*, *J. virol.*, 67:2747 (1993)). Therefore, it is reasonable to expect that compounds that bind with high affinity to this cavity and prevent normal N- and C-helix pairing will be effective HIV-1 inhibitors. In addition, residues in the cavity are highly conserved among diverse HIV-1 isolates. Because of the high structural conservation, drugs targeting this site would have broad activity against diverse HIV-1 isolates, and possibly HIV-2 isolates.

Although this hypothesis is tempting, until now, it had not been demonstrated that these cavity contacts are important for the potency of the C34 inhibitor. In fact, some C-peptides that lack the cavity-binding residues, such as DP178 (C. T. Wild, D. C. Shugars, T. K. Greenwell, C. B. McDanal, T. J. Matthews, *ibid*, 91:9770 (1994); Kilby, J.M. *et al.*, *Nature Med.*, 4:1302 (1998)), are highly effective inhibitors of HIV-1 membrane fusion. These concerns emphasize the need for systematic structure-function analysis to identify determinants of C-peptide activity.

To determine the role of cavity-contacts in inhibitory activity, structure-based mutagenesis was performed on C34. The core of the gp41 ectodomain (Figure 1) was reconstituted with two synthetic peptides called N36 and C34 (M. Lu, P. S. Kim, *J. Biomol. Struct. Dyn.*, 15:465 (1997), D. C. Chan, D. Fass, J. M. Berger, P. S. Kim, *Cell*, 5 89:263 (1997)). Variants of the C34 peptide with single alanine substitutions were synthesized, and the helical content and thermal stability of mutant N36/C34 complexes were quantitated by circular dichroism. As expected, mutation of C34 residues (Met⁶²⁹, Arg⁶³³) that do not contact the N36 coiled-coil had little effect on mean residue ellipticity at 222 nm (a measure of helical content) or stability of N36/C34 complexes 10 (Table 1). However, mutation of the three residues (Trp⁶²⁸ → Ala, Trp⁶³¹ → Ala or Ile⁶³⁵ → Ala) that project into the N36 coiled-coil cavity resulted in N36/C34 complexes with substantially decreased mean ellipticity and stability (Table 1). The greatest destabilization was observed with the mutant Trp⁶³¹ → Ala, which formed N36/C34 complexes with an apparent melting temperature (T_m) of 37°C, compared to 66°C for 15 wildtype. These results demonstrate that C34 residues making hydrophobic contacts with the N36 coiled-coil cavity are important for stabilizing the helical-hairpin structure of the gp41 ectodomain core.

To determine the importance of these residues in the ability of C34 to inhibit membrane fusion, the activity of C34 peptides was tested in HIV-1 viral entry and 20 syncytium assays (Table 1). Mutations (Met⁶²⁹ → Ala and Arg⁶³³ → Ala) that had little effect on the stability of the N36/C34 complex also had little effect on the inhibitory activity of wildtype C34 (IC_{50} ~2.1 nM and ~0.55 nM for viral entry and syncytium formation, respectively). However, mutation of the strictly conserved Trp⁶²⁸ or Trp⁶³¹ to alanine resulted in a substantial decrease in activity of ~5 fold and ~30-fold, respectively 25 (Table 1). Mutation of the less well-conserved Ile⁶³⁵ resulted in only a ~2-fold decrease in inhibitory activity. These results demonstrate for the first time, the C34 residues which make contact with gp41 pocket are important for the inhibitory potency of C34.

The relationship between the potency of mutant C34 peptides and the stability of mutant N36/C34 complexes was clarified by taking advantage of the greatly

destabilizing effect of the Trp⁶³¹ mutation to construct a series of N36/C34 complexes with a gradation of stabilities. The Trp⁶³¹ position was used as a “guest site” and the tryptophan was substituted with natural and artificial amino acids representing a broad range of hydrophobic bulk. In order of increasing hydrophobic bulk, the substitutions

5 were: glycine (Gly), alanine (Ala), L- α -aminobutyric acid (Abu), valine (Val), leucine (Leu), phenylalanine (Phe), the wildtype residue tryptophan (Trp), and L- β -(1-naphthyl) alanine (Nal). This approach resulted in a set of C34 peptides that form N36/C34 complexes with T_ms ranging from 37°C to 66°C. The T_ms and $[\theta]_{222}$ (10³ deg cm²dmol⁻¹) for the N36/C34 variants (with IC₅₀ values (nanomolar) for virus entry and cell fusion,

10 respectively, in parentheses) are: Trp⁶³¹-Gly, 35°C, 17.1 (38 ± 6.1, 25 ± 3.8); Trp⁶³¹-Ala, 37°C, -24.9 (40 ± 4.3, 15 ± 0.8); Trp⁶³¹-Abu, 43°C, -23.2 (16 ± 4.8, 6.9 ± 0.4); Trp⁶³¹-Val, 43°C, -23.9 (13 ± 2.8, 4.5 ± 0.09); Trp⁶³¹-Leu, 50°C, -26.7 (5.3 ± 1.0, 3.2 ± 0.1); Trp⁶³¹-Phe, 59°C, -26.3 (3.6 ± 0.8, 1.6 ± 0.05); wildtype, 66°C, -31.7 (1.5 ± 0.2, 0.55 ± 0.03); Trp⁶³¹-Nal, 62°C, -32.0 (1.4 ± 0.3, 0.79 ± 0.08). The concentration of

15 the Trp⁶³¹-Nal peptide was measured by Nal absorbance using the extinction coefficient $\epsilon = 6900$ at 282 nm (J. Blake, C. H. Li, *J. Med. Chem.*, 18:423-426 (1975)). In HIV-1 infection and syncytium assays, this series of peptides showed potencies that closely correlated with the T_m of the corresponding N36/C34 complex (Figure 2). The potency order of these mutants is wt~Nal>Phe>Leu>Val~Abu>Ala~Gly, in close agreement

20 with the hydrophobic bulk of the substitution and the stability of N36/C34 complexes. There is a striking linear relationship when the IC₅₀ is plotted on a logarithmic scale as a function of the T_m (Figure 2). Since $\Delta G = -RT \ln K$ (ΔG , change in free energy; R, gas constant; T, absolute temperature; and K, equilibrium constant) and ΔT_m (T_{m, wildtype complex} - T_{m, mutant complex}) is proportional to $\Delta(\Delta G)$ ($\Delta G_{\text{wildtype complex}} - \Delta G_{\text{mutant complex}}$) (W. J. Becktel, J.A. Schellman, *Biopolymers*, 26:1859 (1987)), the observed linear relationship strongly

25 suggests that the potency of the C34 variants is directly related to their affinity for the N-helix coiled-coil, as predicted by a dominant-negative mode of inhibition. These results provide strong support for the proposal that the coiled-coil cavity in the gp41 core is an attractive drug target. Conserved residues projecting into the hydrophobic

cavity clearly play a major role in the ability of C34 to inhibit HIV-1 infection, indicating that this inhibitor works by forming a high-affinity complex with the N-helix coiled-coil. Moving beyond traditional peptides, mirror-image phage display techniques (T. N. Schumacher, *et al.*, *Science*, 271:1854 (1996)), selection-reflection aptamer techniques (K.P. Williams *et al.*, *PNAS*, 94:11285 (1997); S. Klußmann *et al.*, *Nat. Biotech.*, 4:1112 (1996); A. Nolte *et al.*, *Nat. Biotech.*, 14:1116 (1996), combinatorial chemistry (A. Borchardt, S. D. Liberles, S. R. Biggar, G.R. Crabtree, S.L. Schreiber, *Chem. Biol.*, 4:961 (1997); J.C. Chabala, *Curr. Opin. Biotechnol.*, 6:632 (1995)) and computational approaches in structure-based drug design (H. Kubinyi, *Curr. Opin. Drug Discov. Develop.*, 1:16 (1998)), can be used to identify D-peptides, peptidomimetics, and small molecules that bind with high affinity to the coiled-coil cavity. The close correlation between N36/C34 stability and C34 inhibitory potency suggests that the effectiveness of such compounds will depend critically on the strength of their cavity-contacts. These results suggest that candidate compounds can be tested for the ability to form a stable complex with N36, thereby providing a basis for rapid, quantitative screens to identify and evaluate potential inhibitors of HIV-1 entry.

Small-molecule inhibitors directed against the cavity of the central coiled-coil target one of the most highly conserved regions of the HIV-1 envelope proteins. The analogous cavity in the SIV gp41 core has an essentially identical structure, with conservation of side chain conformations (V. N. Malashkevich, D. C. Chan, C. T. Chutkowski, P. S. Kim, *Proc. Natl. Acad. Sci. USA*, 95:9134 (1998)). This high degree of structural conservation explains the broad neutralizing activity of C-peptides, which are effective against laboratory-adapted strains as well as primary isolates (C. T. Wild, D. C. Shugars, T. K. Greenwell, C. B. McDanal, T. J. Matthews, *Proc. Natl. Acad. Sci. USA*, 91:9770 (1994), S. Jiang, K. Lin, N. Strick, A.R. Neurath, *Nature*, 365:113 (1993)). Remarkably, SIV C34 peptide is nearly as effective as HIV-1 C34 in inhibiting HIV-1 infection (V. N. Malashkevich, D. C. Chan, C. T. Chutkowski, P. S. Kim, *Proc. Natl. Acad. Sci. USA*, 95:9134 (1998)). In addition, a C-peptide (T649) containing the cavity-binding region is much less susceptible to the evolution of resistant virus (L. T.

Rimsky, D. C. Shugars, T. J. Matthews, *J. Virol.*, 72:986 (1998)) than DP178 (also called T-20), which lacks this region. These observations are evidence that high-affinity ligands targeting the coiled-coil surface, particularly its cavity, will have broad activity against diverse HIV isolates (including HIV-2) and will be less likely to be bypassed by
 5 drug-escape mutants.

These studies on the mechanism of C-peptide action also support the hypothesis that the trimeric hairpin structure of the gp41 core (Chan, D.C. *et al.*, *Cell*, 89:263 (1997); Weissenhorn, W. *et al.*, *Nature*, 387:426 (1997); Tan, K. *et al.*, *Proc. Natl. Acad. Sci. USA*, 94:12303 (1997)) corresponds to the fusion-active state of gp41. The
 10 work described herein shows that the inhibitory potency of C34 depends on its ability to bind to the N-coiled-coil of gp41. Since the hairpin structure of gp41 is extremely stable (with a melting temperature in excess of 90°C) (Lu, M. *et al.*, *Nat. Struct. Biol.*, 2:1075 (1995)), it is unlikely that nanomolar concentrations of C34 can disrupt this structure once it has formed, especially given the high effective concentration of the N-
 15 and C-helices within an intact gp41 molecule. Rather, C-peptides likely act prior to the formation of the gp41 hairpin by binding to a transient pre-hairpin intermediate, in which the central coiled-coil is exposed. Binding of C-peptides to this pre-hairpin intermediate inactivates gp41 and prevents its conversion to the fusion-active hairpin structure (D. C. Chan, P. S. Kim, *Cell*, 93:681 (1998)).

20 As described herein, the pocket on the surface of the N-helix coiled-coil of HIV-1 envelope protein gp41 subunit is a drug target. Similarly, cavities on other pathogens (e.g., HIV-2) which can cause AIDS or on pathogens which cause AIDS-like conditions in nonhuman mammals (e.g., SIV) are also drug targets. As described herein, available methods (e.g., mirror image phage display methods, combinatorial chemistry,
 25 computational approaches and other drug screening and medicinal chemistry methods) can be used to identify peptides, D-peptides, peptidomimetics and small molecules that bind the coiled-coil cavity of HIV-1 (and/or HIV-2) with sufficient affinity to interfere with viral entry into cells and, thus, inhibit viral infection. As further described herein

(Example 3), mirror image phage display has been used to identify D-peptides which bind to a cavity on the surface of the N-helix coiled-coil of HIV-1 gp41.

As a result of the work described herein, screening assays which identify molecules or compounds (agents or drugs) that prevent C34/N36 complex formation and/or disrupt the complex once it has formed are available, as are methods of identifying molecules or compounds (agents or drugs) which bind the N-helix coiled-coil pocket of HIV gp41. Such drugs or agents are useful to inhibit (totally or partially) HIV entry into cells and, thus, infection by HIV.

Methods of screening for compounds or molecules (referred to as drugs or agents) that interfere with formation of a stable complex between C34 and N36 or disrupt a complex between the two and methods of screening for compounds or molecules that bind the N-helix coiled-coil pocket of HIV gp41 are a subject of the present invention.

In one embodiment of a screening method of the present invention, drugs which interfere with formation of a complex between C34 peptide and N36 peptide are identified by combining a candidate drug (a compound or molecule to be assessed for its ability to interfere with formation of a complex between C34 and N36) with C34 and N36, thus forming a test sample, under conditions appropriate for formation of a complex between C34 and N36 and determining whether formation of C34/N36 complex is inhibited (partially or totally) in the test sample. Results of this assessment can be compared with the results of an appropriate control, which is the same combination as the test sample, except that the candidate drug is not present; the control is subjected to the same conditions as is the test sample. If C34/N36 complex is not formed or is formed to a lesser extent in the presence of the candidate drug (in the test sample) than in its absence, the candidate drug is a drug that interferes with formation of a stable complex between C34 and N36. Such a drug is also referred to herein as an inhibitor of C34/N36 complex formation. Inhibition of complex formation can be assessed by determining the extent to which binding of the two members of the complex occurs, such as by means of a fluorescence assay (e.g., FRET), in which C34 and N36

are each labeled by a member of a pair of donor-acceptor molecules or one end of one of the peptides (e.g., the N-terminus of C34) is labeled with one member of such a pair (EDANS) and the natural fluorophore tryptophan, present in the N36 peptide, is the other member of the donor/acceptor pair. Binding of the C34 and N36 is assessed by the extent to which light emission (FRET) occurs from the acceptor model and/or the wavelength spectrum of the light emitted is altered. Prevention of binding by the candidate drug alters the extent to which light is emitted and/or prevents the shift in wavelength that would occur if binding of C34 and N36 occurred. Alternatively, C34 can be labeled with a detectable label, such as a radiolabel (e.g., by synthesizing a variant C34 with a kinase recognition site that can be labeled with a kinase and radioactive ATP). The radiolabeled C34 and the candidate drug are combined with N36 immobilized to, for example, a solid surface (e.g., a bead or a plastic well), thus producing a test sample. The extent to which binding of labeled C34 with immobilized N36 occurs is determined and compared with the extent to which binding of labeled C34 to immobilized N36 occurs under the same conditions to which the test sample is subjected, but in the absence of the candidate drug (in a control sample). Typically, this assessment is carried out after the sample has been maintained for sufficient time and under appropriate conditions for C34/N36 binding to occur and a subsequent wash to remove any unbound C34 and candidate drug. If binding occurs in the test sample to a lesser extent than in the control sample, as evidenced by less radiolabel bound to the immobilized N36 in the test sample than in the control sample, the candidate drug is an inhibitor of binding of C34 and N36. Alternatively, the label or tag on C34 can be a member of a binding pair, the other member of which is used to detect binding to N36. For example, C34 can be tagged with biotin (through standard solid-state peptide synthesis, for example) and combined with N36, which can be in solution or bound to a solid surface, such as a bead, well or flat/planar surface, along with the candidate drug (test sample) or in the absence of the candidate drug (control sample). Binding of C34 to N36 is assessed by detecting the presence of biotin associated with N36, such as through the use of labeled streptavidin (e.g., streptavidin - HRP, streptavidin - AP or

iodinated streptavidin), which binds the biotin on C34 and is then itself detected through its label. If binding occurs less in the presence of the candidate drug (in the test sample) than in the absence of the candidate drug (in the control sample), as indicated by the presence of less biotin detected on N36 in the test sample than in the control sample, the

5 candidate drug is an inhibitor of C34/N36 binding. The candidate drugs can be obtained, for example, from a library of synthetic organic compounds or random peptide sequences, which can be generated synthetically or through recombinant technology.

In a similar fashion, the ability of a candidate drug to disrupt C34/N36 binding can be assessed, to identify inhibitors of C34/N36 and, thus, of HIV infection. In this

10 embodiment, preformed C34/N36 complex is combined with a candidate drug, which is to be assessed for its ability to disrupt the complex, thus producing a test sample. The control sample is the same as the test sample, except that the control sample does not contain the candidate drug; it is treated in the same manner as the test sample. If

15 C34/N36 binding is disrupted in the presence of the candidate drug and not in the control sample or if disruption of the complex occurs to a greater extent in the test sample than in the control sample, the candidate drug is an inhibitor (disrupter) of C34/N36. Detection of disruption of binding can be carried out as described above for detection of/prevention of/interference with binding of C34/N36 (e.g., by FRET or a fluorescence assay, by detecting a radiolabel or other detectable label, such as biotin.)

20 Results described herein demonstrate that hybrids (i.e., fusion proteins) can be made between a trimeric version of the coiled-coil region of a protein (such as GCN4) and the N-helix coiled-coil of HIV gp41, and that such hybrids are trimeric (i.e., not aggregated) and 100% helical. Results described herein also clearly indicate that such fusion proteins do not disrupt or alter the structure of the N-peptide large cavity (i.e.,

25 hydrophobic pocket), which is essentially the same in IQN17 (ligand-free and in complex with D10pep1; see Example 5) as it is in N36 (i.e., in complex with C34; Chan D.C. *et al. Cell*, 89, 263 (1997)).

Figures 5A, 5B and 6 present results of assessment of peptides described herein. In Figure 5A-5B, the IQN17 crystal structure is shown to be a continuous, three-

stranded coiled-coil; the 17 residues derived from HIV gp41 form a hydrophobic pocket very similar to that found in the crystal structure of gp41. As shown, D10pep1 is bound to this pocket and the residues of D10pep1 that correspond to the conserved residues (leucine, tryptophan, tryptophan) found in all of the D-peptide inhibitors described

5 herein are packed into this pocket, clearly indicating that other D-peptide inhibitors which comprise these conserved residues would bind to IQN17 in the same manner. Figure 6 shows results of syncytia assays carried out according to the method described by Chan *et al.* (Chan, D. C. *et al. Proc. Natl. Acad. Sci.*, 95: 15613-15617 (1998)). In the experiments whose results are represented in Figure 6, D-peptides identified as

10 described herein were used. In each instance, a blocking group (e.g., an acetyl group) was present at the N terminus and a CONH₂ (amide) was present at the C-terminus. Results of these assays showed a range of IC₅₀ concentrations, where IC₅₀ is the concentration at which one half of the number of syncytia are observed, compared to the control, in which no peptide is included. For example, D10pep5 with two lysines at the

15 N-terminus has an IC₅₀ of approximately 6μM.

In another embodiment, the invention relates to a method of identifying a drug that binds the N-helix coiled-coil cavity of HIV gp41. Here, too, the assay is based on assessing loss or decrease in binding, but unlike the C34/N36 complex assay described above, which is a more general assay in that it covers or detects interaction with any

20 portion of the groove formed by the N-helical region of HIV gp41, this embodiment focuses on the HIV gp41 hydrophobic pocket (the N-helix coiled-coil cavity). In this embodiment, the method comprises combining a candidate drug to be assessed for its ability to bind the N-helix coiled-coil cavity of HIV gp41 with a fusion protein that comprises a trimeric version of the coiled-coil region of a protein and a sufficient

25 portion of the N-peptide of HIV gp41 to include the HIV gp41 cavity, under conditions appropriate for presentation of the HIV gp41 cavity for binding by a peptide or other molecule and determining (e.g., in a high throughput screen) whether the candidate drug binds the fusion protein. If binding occurs, the candidate drug is a “hit” that may be a drug that binds the N-helix coiled-coil cavity of HIV gp41. If binding occurs, the

candidate drug has bound the N-helix coiled coil and it can be determined if it binds to the coiled-coil cavity. Such "hits" can then be screened in secondary assays, such as the cell/cell fusion assay and HIV infectivity assay to determine if the candidate drug is a drug. Alternatively, or in addition, such "hits" can be assessed further by use of a

5 counterscreen with other fusion proteins (or peptides), to which pocket-binding molecules will not bind. For example, GCN4-pIQI (with the same three surface mutations as in IQN17) or a version of IQN17 with a point mutation in the hydrophobic pocket, IQN17(G39W), in which glycine 39 is mutated to tryptophan, resulting in a large protrusion into the pocket, can be used in a counterscreen. In this example, a

10 candidate drug that binds to IQN17 but not to GCN4-pIQI (with the same three surface mutations as in IQN17) or IQN17(G39W) is a drug that binds the N-helix coiled-coil cavity of HIV gp41.

In a further embodiment, a competitive assay is carried out. In this embodiment, a peptide or protein that binds the N-helix coiled-coil cavity of HIV gp41 is combined

15 with the candidate drug and the fusion protein and whether the candidate drug binds the HIV gp41 cavity is determined in the presence of the peptide that binds the N-helix coiled cavity of HIV gp41. If the candidate drug binds the fusion protein, it is a drug that binds the HIV gp41 cavity. For example, a fusion protein which comprises a trimeric version of the coiled-coil region of GCN4 and the C-terminus of the N peptide

20 of HIV gp41 that includes the N-helix coiled-coil cavity (IQN17) is combined with a "reference" D-peptide (e.g., any of the D-peptides described herein or variants thereof) that binds the N-helix coiled-coil cavity and a candidate drug to be assessed for its ability to bind the N-helix coiled-coil cavity of HIV gp41, thus producing a test sample, which is maintained under conditions appropriate for binding of the D-peptide to bind to

25 the cavity. A control sample, which includes the same components as the test sample, except for the candidate drug, and is handled in the same manner as the test sample, is also assessed. In both samples, binding of the reference D-peptide is assessed. If binding of the reference D-peptide occurs to a lesser extent in the presence of the candidate drug (in the test sample) than in its absence (in the control sample), the

candidate drug is a drug that binds the N-helix coiled-coil cavity of HIV gp41.

Detection of binding is assessed, for example, in a similar manner as described above for the C34/N36 embodiment of the invention. For example, the D-peptide is labeled with a detectable label, such as a radiolabel or a first member of a binding pair (e.g.,
5 biotin), and the extent to which the N-helix coiled-coil cavity bears the label (after the samples have been maintained under conditions appropriate for binding of the reference D-peptide to the cavity) is determined. In the case in which radiolabeling is used, the extent to which the fusion protein bears the radiolabel is assessed in the test sample and compared with the extent to which the fusion protein bears the radiolabel in the control
10 sample. If the detectable label is a first member of a binding pair (e.g. biotin), the second member of the pair (a binding partner) is added to the samples in order to detect the extent to which the fusion protein is bound by the reference D-peptide. This can be done directly or indirectly (e.g., by adding a molecule, such as an antibody or other moiety which binds the second member of the binding pair). Less of the label will be
15 present on the fusion protein (N-helix coiled-coil cavity) if the candidate drug has inhibited (totally or partially) binding of the D-peptide to the cavity. If binding occurs to a lesser extent in the test sample (in the presence of the candidate drug) than in the control sample (in the absence of the candidate drug), then the candidate drug is a drug that binds the N-helix coiled-coil cavity of HIV gp41.

20 IQN17, or a variant thereof, in the D-enantiomer, is useful to identify molecules or compounds which are members of a library or collection and bind the N-helix coiled-coil of gp41. For example, a library or collection of molecules or compounds, such as a phage display library, can be screened with IQN17 in the D-enantiomer to identify members that bind the pocket. This has been carried out successfully, as described
25 herein. The mirror image of IQN17, or a variant thereof, is used as the target molecule. As used herein, the terms "D-enantiomer of a polypeptide" and "D-peptide" refer to the exact mirror image of the molecule in the natural handedness. Thus, for amino acid residues that contain a second chiral center, such as Ile and Thr, the exact mirror image of the naturally-occurring amino acid residue is used to create the D version of the

polypeptide. Also as used herein, the terms “D-amino acids” and “L-amino acids” are both meant to include the non-chiral amino acid glycine. D-IQN17 can be immobilized to a solid surface, such as by addition of one member of a binding pair (e.g., biotin) to it and addition of the other member of the pair (e.g., streptavidin) to the solid surface.

- 5 Binding of the two members results in immobilization of D-IQN17 on the solid surface, such as for phage panning. A linker which is an enzyme recognition site (e.g., an amino acid linker such as Gly-Lys-Gly, in which an L-lysine residue is used) can be placed between the D-IQN17 sequence and the binding pair member (between the biotin and D-IQN17) to provide an enzyme recognition site (here, a trypsin recognition site), so
- 10 that bound phage can be eluted by a trypsin digestion, rather than by non-specific elution, such as acid addition. The phage display library can be a library of L-amino acid peptides of any appropriate length fused to an appropriate phage gene. In one embodiment, it is a phage display library of L-amino acid peptides fused to the gIII gene of M13 phage. The peptides, in one embodiment, comprise 10 randomly encoded
- 15 amino acid residues flanked by either a cysteine or a serine on both sides. Typically, several rounds of panning are carried out. D-IQN17-specific binding phage are identified. Phage that bind only the gp41 region of D-IQN17 can be identified by post-panning assessment, such as by screening against wells that lack the antigen and then further testing against a panel of molecules. For example, specific pocket-binding
- 20 phage include those that bind D-IQN17 but not D-GCN4-pIQI (with the same three surface mutations as in IQN17) or a version of D-IQN17 with a point mutation in the hydrophobic pocket, D-IQN17(G39W), in which glycine 39 is mutated to tryptophan, resulting in a large protrusion into the pocket. D-peptides identified in this manner can be assessed for their ability to inhibit HIV gp41, using known assays, such as the
- 25 cell/cell fusion assay and HIV infectivity assay. The mirror-image phage display method described herein has demonstrated the value of IQN17 and IQN17(G39W), and their D-enantiomers in identifying inhibitors of HIV-1 entry that bind the gp41 pocket. Of nine specific pocket-binding phage sequences identified (phage that bind to D-IQN17 but not to D-IQN17(G39W), eight contain a consensus EWXWL sequence and

inhibit HIV-1 gp41-induced syncytia formation when tested as D-peptides. The ninth peptide was toxic to cells and was not investigated further.

The D-versions of IQN17 and IQN17(G39W) can be used in a similar manner with other biologically encoded libraries, to discover other pocket-binding molecules that are not subject to enzymatic degradation by natural enzymes. For example, other phage-display libraries can be used to identify new D-peptide inhibitors (e.g., with a different number of residues between the flanking Cys residues and/or with randomly encoded amino acid residues outside the regions flanked by cysteine residues and/or with more than two cysteine residues). Strategies for encoding peptide libraries without phage (e.g., in which the encoding mRNA is attached to the peptide) can be used to identify D-peptide inhibitors. RNA or DNA libraries can be used (e.g., with SELEX methods) to identify L-ribose- or L-deoxyribose-based RNA or DNA aptamers, respectively, that bind to the hydrophobic pocket and are not substrates for natural nucleases (see e.g., Williams *et al.*, *PNAS*, 74:11285 (1997)).

Although the versions of IQN17 and IQN17(G39W) of natural L-handedness can also be used in similar manner with biologically encoded libraries, the most likely applications will be with other, non-biologically encoded libraries. For example, chemical combinatorial libraries on beads (of the one-bead, one-compound variety) can be screened with labeled IQN17 (e.g., radioactive or with a chromophore) to identify beads containing molecules that bind to IQN17. In this example, IQN17(G39W) can be used as a counterscreen to determine if the molecules on the bead bind to the pocket of IQN17. (If they bind to IQN17(G39W), then they are not likely to be pocket-binding molecules.) As another example, beads to which IQN17 had been previously attached can be incubated with a mixture of potential pocket-binding molecules (e.g., a mixture of chemicals, or a natural product extract). IQN17 (bound to the beads) can then be separated from the mixture, washed, and then subjected to conditions (e.g., organic solvent, low pH, high temperature) that elute molecules bound to the IQN17 on the beads. The eluted molecules (i.e., potential pocket-binding molecules) could be identified by analytical chemistry methods (e.g., HPLC, mass spectrometry). A

counterscreen with IQN17(G39W) is useful to help to identify true pocket-binding molecules.

Drugs identified by the methods described above are then further tested for their ability to inhibit (totally or partially) HIV gp41 function (membrane fusion) and, thus entry into cells, using further *in vitro* assays, such as the syncytium assays and/or infectivity assays described herein or others known to those of skill in the art, and/or *in vivo* assays in appropriate animal models or in humans.

One embodiment of the present invention is a method of identifying a drug that binds the N-helix coiled-coil of HIV gp41, particularly the N-helix coiled-coil pocket.

10 The method comprises combining a candidate drug to be assessed for its ability to bind the N-helix coiled-coil pocket of HIV gp41 and peptide which comprises a soluble, trimeric coiled-coil and a sufficient portion of the N-peptide of HIV gp41 to include the HIV gp41 pocket, under conditions appropriate for presentation of the HIV gp41 pocket for binding by a molecule or compound (e.g., a drug) and determining whether the

15 candidate drug binds the HIV gp41 pocket. If binding of the candidate drug with the HIV gp41 pocket occurs, the candidate drug is a drug which binds the N-helix coiled-coil pocket of HIV gp41. Optionally, binding of the candidate drug can be assessed in the assay as described above, except that a peptide that binds the N-helix coiled-coil pocket (a peptide previously identified as one which binds the pocket) is combined with

20 the candidate drug and the peptide. In this competitive assay, binding of the candidate drug to the N-helix coiled-coil pocket is assessed in the presence of a known binding moiety (a molecule or compound which binds the pocket). If binding of the candidate drug occurs in the presence of the known binding moiety, the candidate drug is a drug which binds the N-helix coiled-coil pocket with sufficient affinity to successfully

25 compete with the known binding moiety. The fusion protein used in this embodiment comprises a soluble, trimeric version of a coiled-coil, such as a soluble, trimeric version of the coiled-coil region of a protein (e.g., a non-HIV protein, such as that of GCN4 or GCN4-pIQI, although an HIV protein can be used) and a sufficient portion of the N-peptide of HIV gp41 to include the HIV gp41 cavity. For example, this portion can

comprise SEQ ID NO.: 20 or a sufficient portion to comprise the cavity and, when present in an appropriate fusion protein or other soluble model, present the cavity in such a manner that it is available for binding. Alternatively, a variant of the HIV gp41 sequence present herein, a sequence from another strain of the human virus (e.g., HIV-2) or a sequence from another species (e.g., SIV, feline immunodeficiency virus, Visna virus (M. Singh *et al.*, *J. Mol. Biol.*, 290:1031 (1999)) can be used in the fusion protein or soluble model. The fusion protein can comprise a soluble, trimeric version of the coiled-coil of any protein, provided that when it is in the fusion protein with the HIV component, the HIV cavity is presented in such a manner that it is available for binding. It can be, for example, that of GCN4-pI_QI, GCN4-pII, Moloney Murine Leukemia Virus (Mo-MLV) or the ABC heterotrimer. In one embodiment, the fusion protein is IQN17 in the D- form. In another embodiment, the fusion protein is IQN17 in the natural L-handedness.

In the competitive assay format, any peptide known to bind the N-helix coiled-coil cavity can be used as the known binding moiety. For example, any of the peptides described herein (SEQ ID NOS.: 3-12, 15, 17-19, 23, 24) or a variant or portion thereof can be used. Also, any non-peptide pocket-binding molecule can be used in the competitive assay format. The competitive assay can be performed in solution, on a bead, or on a solid surface.

In one embodiment, the candidate drug is detectably labeled and binding of the candidate drug to the HIV gp41 N-helix coiled-coil is determined by detecting the presence of the detectable label on the HIV gp41 N-helix coiled-coil (as a result of binding of the labeled candidate drug to the N-helix coiled-coil). Detection of the label on the helix coiled-coil pocket of the soluble model is indicative of binding of the candidate drug to the N-helix coiled-coil pocket and demonstrates that the candidate drug is a drug which binds the N-helix coiled-coil pocket. If the labeled candidate drug is detected on the fusion protein, the candidate drug is a drug which binds the N-helix coiled-coil cavity.

In another embodiment of the method of identifying a drug that binds the N-helix coiled-coil pocket of the HIV gp41, a soluble model that presents the pocket in such a manner that it is available for binding by a drug is combined with a candidate drug and whether binding of the candidate drug with the N-helix coiled-coil of the
5 soluble model occurs is determined. If binding occurs, the candidate drug is a drug which binds the pocket. Here, too, a competitive assay format can be used. The components of the competition assay (e.g., IQN17 and a D-peptide) can be labeled, with any of a variety of detectable labels, including fluorophore/quencher combinations. The candidate drug can be labeled, as described above, with any of a variety of detectable
10 labels. The components of the soluble model (fusion protein) used in this embodiment and the competing moiety which is used in a competitive assay format can also be as described above.

The present invention also relates to a method of producing a drug that binds the N-helix coiled-coil pocket of HIV gp41. In one embodiment, the method is carried out
15 as follows: A soluble model that presents the N-helix coiled-coil pocket of HIV gp41 or a fusion protein which comprises a soluble, trimeric coiled-coil (e.g., of a protein, such as a non-HIV protein, such as GCN4-pIQI, GCN4-pII, Mo-MLV, ABC heterotrimer or an HIV protein) is combined with a candidate drug to be assessed for its ability to bind the N-helix coiled-coil pocket of HIV gp41 and inhibit entry into cells, under conditions
20 appropriate for presentation of the HIV gp41 pocket for binding by a drug. Whether the candidate drug binds the HIV gp41 pocket is determined, wherein if binding of the candidate drug to the N-helix coiled-coil pocket of HIV gp41 occurs, the candidate drug is a drug which binds the N-helix coiled-coil cavity of HIV gp41. In this embodiment, the fusion protein comprises a soluble, trimeric coiled-coil (e.g., of a protein such as a
25 non-HIV protein, such as a soluble, trimeric coiled coil of GCN4, GCN4-pIQI, GCN4-pII, Mo-MLV, ABC heterotrimer or an HIV protein) and a sufficient portion of the N-peptide of HIV gp41 to include the HIV gp41 N-helix coiled-coil pocket (e.g., all or a portion of SEQ ID NO.: 20, a variant or modification thereof or a sequence from another strain or species). IQN17, described herein, can be used in this method; the D

enantiomer of IQN17 can also be used (e.g., in mirror-image phage applications). The ability of the drug produced to inhibit HIV entry into cells is assessed, for example, in a syncytium assay and/or an infectivity assay, as described herein. It can be further assessed in an appropriate animal model or in humans.

- 5 The invention also relates to a method of producing a drug that binds the N-helix coiled-coil pocket of HIV gp41. The method comprises: producing or obtaining a soluble model of the N-helix coiled-coil pocket of HIV gp41 (e.g., a fusion protein as described herein and particularly IQN17 or a variant thereof); combining a candidate drug (a molecule or compound) to be assessed for its ability to bind the N-helix coiled-coil pocket of HIV gp41 and the soluble model of the N-helix coiled-coil pocket of HIV gp41 and determining whether the candidate drug binds the N-helix coiled-coil pocket of HIV gp41. If the candidate drug binds the N-helix coiled-coil pocket of HIV gp41, the candidate drug is a drug which binds the N-helix coiled-coil pocket of HIV gp41; as a result, a drug which binds the N-helix coiled-coil cavity of HIV gp41 is produced.
- 10 The fusion protein used in this embodiment is described herein and can be, for example, IQN17, the D enantiomer of IQN17, or variants thereof. Alternatively, a drug that binds the N-helix coiled-coil pocket of HIV gp41 and inhibits entry of HIV into cells can be produced by a method comprising: producing or obtaining a soluble model of the N-helix coiled-coil pocket of HIV gp41, as described herein; combining the soluble model and a candidate drug to be assessed for its ability to bind the N-helix coiled-coil pocket of HIV gp41; determining whether the candidate drug binds the N-helix coiled-coil pocket of the soluble model (fusion protein), wherein if binding occurs, the candidate drug is a drug which binds the N-helix coiled-coil of HIV gp41; and assessing the ability of the drug which binds the N-helix coiled-coil to inhibit HIV entry into cells, wherein if the drug inhibits HIV entry into cells, it is a drug which binds the N-helix coiled-coil pocket of HIV gp41 and inhibits HIV entry into cells. Its ability to inhibit HIV entry into cells can be assessed *in vitro* (e.g., in a syncytium assay, an infectivity assay) or *in vivo* (e.g. in an appropriate animal model or in humans). The soluble model can be a peptide which comprises a soluble, trimeric coiled-coil, such as that of a
- 15
- 20
- 25

protein (e.g., GCN4-pI_QI) and a sufficient portion of the N-peptide of HIV gp41 to include the HIV gp41 pocket.

Drugs identified or produced by the methods described herein, as well as by other methods, which bind the N-helix coiled-coil pocket of HIV gp41 and inhibit HIV entry into cells are also the subject of this invention.

Drugs identified or produced by the methods described herein, as well as by other methods, which bind to more than one N-helix coiled-coil pocket of HIV gp41 and inhibit HIV entry into cells are also the subject of this invention. Such drugs can be obtained, for example, by linking two or more pocket-binding molecules (drugs) via an appropriate linker (e.g., a linker of amino acid residues or other chemical moieties) to increase the effectiveness of inhibition. The pocket-binding molecules that are linked can be the same or different. Drugs identified or produced by the methods described herein or by other methods which bind to the N-helix coiled-coil pocket of HIV gp41, in addition to binding to HIV gp120, CD4, CCR5, CXCR4, or a non-pocket region of HIV gp41 are also the subject of this invention.

Drugs which inhibit HIV gp41 can also be designed or improved with reference to the X-ray crystal structure of the complex between IQN17 and a D-peptide which binds the N-helix coiled-coil cavity presented by IQN17, such as with reference to the X-ray structure of the complex between IQN17 and D10pep1, presented herein. Alternatively, or in addition, drugs which inhibit HIV gp41 can also be designed or improved with reference to the X-ray crystal structure of free IQN17, presented herein.

Compounds and molecules (drugs) identified as described herein inhibit (partially or totally) entry of HIV into cells, and thus are useful therapeutically in uninfected individuals (humans) and infected individuals (e.g., to prevent or reduce infection in an uninfected individual, to reduce or prevent further infection in an infected individual) and as research reagents both to study the mechanism of gp41-induced membrane fusion and to assess the rate of viral clearance by an individual and as reagents to discover or develop other compounds and molecules (drugs) that inhibit entry of HIV into cells. D-peptides described herein (e.g., D10pep5, D10pep1) have

been shown, using the infectivity assay described herein, to inhibit infection of cells. Other D-peptides can be similarly assessed for their ability to inhibit infectivity.

The drugs can be administered by a variety of route(s), such as orally, nasally, intraperitoneally, intramuscularly, vaginally or rectally. In each embodiment, the drug is provided in an appropriate carrier or pharmaceutical composition. For example, a cavity-binding drug can be administered in an appropriate buffer, saline, water, gel, foam, cream or other appropriate carrier. A pharmaceutical composition comprising the drug and, generally, an appropriate carrier and optional components, such as stabilizers, absorption or uptake enhancers, flavorings and/or emulsifying agents, can be formulated and administered in therapeutically effective dose(s) to an individual (uninfected or infected with HIV). In one embodiment, drugs which bind the N-helix coiled-coil of gp41 (e.g., those described herein, DP178 (C. T. Wild, D. C. Shugars, T. K. Greenwell, C. B. McDanal, T. J. Matthews, *ibid*, 91:9770 (1994)), T649 which corresponds to residues 117-152 of HIV-1 gp41 (HXB2 strain) and is acetylated at the amino terminus and amidated at the carboxy terminus) (L. T. Rimsky, D. C. Shugars, T. J. Matthews, *J. Virol.*, 72:986 (1998), are administered (or applied) as microbicidal agents and interfere with viral entry into cells. For example, a drug or drugs which bind(s) the HIV cavity can be included in a composition which is applied to or contacted with a mucosal surface, such as the vaginal, rectal or oral mucosa. The composition comprises, in addition to the drug, a carrier or base (e.g., a cream, foam, gel, other substance sufficiently viscous to retain the drug, water, buffer) appropriate for application to a mucosal surface or to the surface of a contraceptive device (e.g., condom, cervical cap, diaphragm). The drug can be applied to a mucosal surface, such as by application of a foam, gel, cream, water or other carrier containing the drug. Alternatively, it can be applied by means of a vaginal or rectal suppository which is a carrier or base which contains the drug or drugs and is made of a material which releases or delivers the drug (e.g., by degradation, dissolution, other means of release) under the conditions of use (e.g., vaginal or rectal temperature, pH, moisture conditions). Such compositions can also be administered orally (e.g., swallowed in capsule, pill, liquid or other form) and

pass into an individual's blood stream. In all embodiments, controlled or time release (gradual release, release at a particular time after administration or insertion) of the drug can be effected by, for example, incorporating the drug into a composition which releases the drug gradually or after a defined period of time. Alternatively, the drug can

5 be incorporated into a composition which releases the drug immediately or soon after its administration or application (e.g., into the vagina, mouth or rectum). Combined release (e.g., release of some of the drug immediately or soon after insertion, and over time or at a particular time after insertion) can also be effective (e.g., by producing a composition which is comprised of two or more materials: one from which release or

10 delivery occurs immediately or soon after insertion and/or one from which release or delivery is gradual and/or one from which release occurs after a specified period). For example, a drug or drugs which bind the HIV cavity can be incorporated into a sustained release composition such as that taught in U.S. Patent 4,707,362. The cream, foam, gel or suppository can be one also used for birth control purposes (e.g., containing a

15 spermicide or other contraceptive agent), although that is not necessary (e.g., it can be used solely to deliver the anti-HIV drug, alone or in combination with another non-contraceptive agent, such as an antibacterial or antifungal drug or a lubricating agent). An anti-HIV drug of the present invention can also be administered to an individual through the use of a contraceptive device (e.g., condom, cervical cap, diaphragm) which

20 is coated with or has incorporated therein in a manner which permits release under conditions of use a drug or drugs which bind the HIV gp41 N-helix coiled coil. Release of the drug(s) can occur immediately, gradually or at a specified time, as described above. As a result, they make contact with and bind HIV and reduce or prevent viral entry into cells.

25 In another embodiment, a drug which interferes with HIV entry into cells by a mechanism other than binding to the gp41 N-helix coiled-coil cavity (e.g., a drug which interferes with viral entry by interfering with gp120 binding at the CD4 stage) is administered or applied to a mucosal surface as described above for drugs which bind to the gp41 N-helix coiled coil.

Fusion proteins of the present invention comprise a soluble, trimeric form or version of a coiled-coil, such as a soluble, trimeric form or version of a coiled-coil region of a protein (of non-HIV origin or of HIV origin) and a sufficient portion of the C-terminal end of the N peptide of HIV gp41 to include (comprise) the HIV coiled-coil cavity or hydrophobic pocket (the pocket-comprising residues of the N-peptide). The N peptide of HIV gp41 can be that of HIV-1, HIV-2, another HIV strain or a strain from another species (e.g., simian immunodeficiency virus (SIV), feline immunodeficiency virus or Visna virus). For example, HIV-2 sequence LLRLTVWGTKNLQARVT (SEQ ID NO: 26), SIV sequence LLRLTVWGTKNLQTRVT (SEQ ID NO: 27) or a sequence comprising invariant residues in HIV-1, HIV-2 and SIV (represented LLXLTVWGXXXLQXRX (SEQ ID NO: 42), wherein amino acid residues L, T, V, W, G, K, Q, and R are the single letter code used for amino acid residues and X can be any amino acid residue). Also the subject of this invention is a soluble trimeric model of the HIV gp41 hydrophobic pocket, which can be a D-peptide or an L-peptide and comprises a soluble trimeric coiled coil and a sufficient portion of the N peptide region of HIV gp41 to comprise the amino acid residues which form the pocket of the N-helix coiled-coil region of HIV gp41. The D- or L-peptide can comprise as the soluble, trimeric coiled coil the coiled coil of GCN4-pI_QI, of GCN4-pII, of Moloney Murine Leukemia Virus or of the ABC heterotrimer. The component which is a sufficient portion of the N peptide of HIV gp41 to comprise the amino acid residues of the pocket can comprise, for example: LLQLTVWGIKQLQARIL of HIV-1 (SEQ ID NO: 20); LLRLTVWGTKNLQARVT of HIV-2 (SEQ ID NO: 26); LLRLTVWGTKNLQTRVT of SIV (SEQ ID NO: 27) or the invariant residues of these, which are: LLXLTVWGXXXLQXRX (SEQ ID NO: 42).

One embodiment of the instant invention are fusion proteins between a trimeric version of the coiled-coil region of a protein (such as GCN4-pI_QI) and the N-helix coiled-coil of HIV gp41 that include all, part or none of the N-helix cavity. That is, a fusion protein of the present invention can comprise a trimeric form of the coiled-coil region of GCN4-pI_QI and a portion of the N-peptide of HIV-1 gp41, wherein the portion

of the N-peptide of gp41 comprises part, or all, or none of the N-helix cavity of HIV-1 gp41. For example, a fusion protein can be made that contains residues from GCN4-pI_QI and residues from N36. The fusion protein, denoted IQN24n, contains 29 residues of GCN4-pI_QI, including three mutations for increased solubility, and 24 residues from the N-terminal end of N36 (SGIVQQQNNLLRAIEAQQHLLQLT) (SEQ ID NO 21); for recombinant expression in *E. coli*, an extra Met residue is included at the N-terminus. For example, a fusion protein can comprise a portion of the N-peptide of HIV gp41 comprising the amino acid sequence of (SEQ ID.: 21). The sequence of IQN24n is:

10 MRMKQIEDKIEEIESKQKKIENEIARIKKLISGIVQQQNNLLRAIEAQQHLLQLT (SEQ ID.: 22). This fusion protein can be made by a variety of methods, including chemical synthesis or recombinant DNA methods or by recombinant expression in *E. coli*, in which case the N- and C-termini are not blocked. Because the superhelix parameters of the GCN4-pI_QI coiled coil are nearly identical to the HIV gp41 N-helix coiled coil, the resulting fusion protein molecule (IQN24n) is predicted to form a long trimeric coiled coil, which presents part of the gp41 N-helix coiled coil as a trimer (not aggregated).

An alternative embodiment of the instant invention provides a method of eliciting an immune response in an individual. The strategy used to create a soluble, trimeric model for part of the gp41 N-terminal region coiled coil is also helpful to develop HIV vaccine candidates. One goal for a potential HIV vaccine is to elicit a neutralizing antibody response that binds to the “pre-hairpin” intermediate of the HIV-1 gp120/gp41 envelope protein complex. In this transient form, the N-helix region of gp41 is exposed, but the C-helix region is not. Although it seems reasonable to use an N-peptide (such as N36, N51 or DP-107) as an immunogen to elicit an antibody response against the N-helix region of gp41, the isolated N-peptides are aggregated and do not properly present the gp41 N-helix coiled-coil trimer. Accordingly, the same strategy described herein to solve this problem for the gp41 hydrophobic pocket can be applied towards the development of soluble, trimeric models of the gp41 N-helix coiled-

coil region, in general. Such trimeric models (including IQN 17, but also including, for example, peptides that do not contain the pocket residues of gp41) can be used as immunogens to elicit an antibody response to the pre-hairpin intermediate, thereby inhibiting HIV-1 infection. For example, an individual to be immunized can be

5 administered a fusion protein comprising a trimeric form of a coiled-coil region of a protein and a portion of an N-peptide from HIV-1 gp41, wherein the portion from gp41 comprises part of, all of, or none of the N-helix coiled-coil cavity in a pharmaceutically acceptable carrier. For example, IQN24n can be used, either alone or in combination with other materials, in a vaccine, which will elicit the production of antibodies that

10 bind to the coiled coil in the individual to whom it is administered (the vaccinee), and thereby offer protection against infection and/or disease. IQN24n can also be used to identify (from humans, other animals or antibody libraries) and/or raise antibodies (monoclonal and/or polyclonal) that bind to the N-helix coiled coil. This provides the basis for a diagnostic method in which IQN24n (or IQN17 or other soluble trimeric

15 model) is used to assess the presence/absence/level of antibodies that bind the N-helix coiled coil in a biological sample (e.g., blood).

Any of a wide variety of variations can be made in the GCN4-pI_QI component of fusion proteins described herein (e.g., IQN17 or IQN24n) and used in the method, provided that these changes do not alter the trimeric state of the coiled-coil. Changes

20 can also be made in the amino acid composition of the fusion protein component which is the portion from the HIV gp41 N36 peptide, to produce variants (e.g., variants of IQN17 or IQN24n). There is no limit to the number or types of amino acid residue changes possible, provided that the trimeric state of the coiled-coil and the structure of the surface of the fusion protein corresponding to the N-peptide coiled coil of HIV gp41

25 are maintained. The fusion protein component which is the portion of the HIV gp41 N-peptide can include all, part, or none of the N-helix cavity. For example, other parts of N51, N36, DP-107, or other regions of the HIV gp41 N-helix region can be fused to GCN4-pI_QI (or another trimeric version of the coiled-coil region of a protein) to generate trimeric (not aggregated) helical coiled-coil fusion proteins and used in the

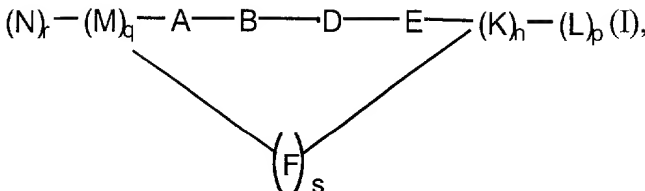
method. There is no limit to the number or types of fusion proteins that can be designed and generated, provided that the trimeric state of the coiled-coil and the structure of the surface of the fusion protein corresponding to the N-peptide coiled coil of HIV gp41 are maintained. Such fusion proteins can be designed and generated using methods known
5 to those of skill in the art, such as evaluating heptad-repeat positions or superhelix parameters of coiled coils.

Described herein are peptides, which can be D-peptides or L-peptides, which bind to a cavity on the surface of the N-helix coiled-coil of HIV envelope glycoprotein gp41 (e.g., HIV-1, HIV-2). Such peptides can be of any length, provided that they are of
10 sufficient length to bind the cavity in such a manner that they interfere with the interaction of the N-helix coiled-coil cavity and amino acid residues of the C-peptide region of HIV gp41 and prevent HIV entry into the cells. For example, D- or L-peptides comprise at least two amino acid residues and generally will be from about two to about 21 amino acid residues. That is, they can comprise any number of amino acid residues
15 from about two to about 21. The amino acid residues can be naturally occurring or non-naturally occurring or modified, as described below. The peptides can be linear or circular.

Examples of D-peptides, identified as described herein, are shown in Figure 3. Because of library design, each peptide, in addition to the amino acid residues shown, is
20 flanked by GA on the N-terminus and AA on the C-terminus. N-terminal lysine residues were added to improve water solubility.

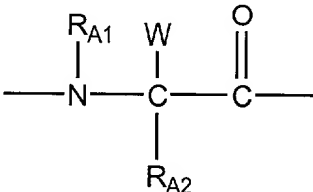
In one embodiment, the present invention provides compounds which inhibit the binding of the N-helix coiled coil to the C-helix of HIV-1 gp41 envelope protein. Such compounds are of use in a method of treating a patient infected by, or potentially subject
25 to infection by, HIV. These compounds are also of use in a method of assessing the ability of a second compound to bind to the N-helix coiled coil cavity.

In one embodiment, the compounds which inhibit the binding of the N-helix coiled coil to the C-helix of HIV-1 gp41 envelope protein are of Formula I,



5

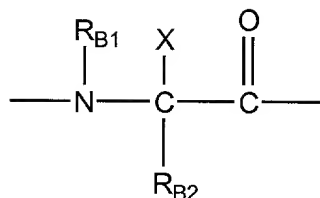
10



15

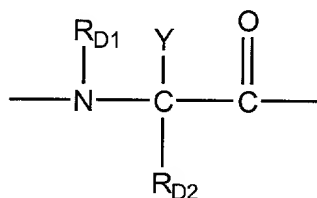
B is a glycyl residue or D-amino acid or N-substituted glycyl residue of the formula

-47-



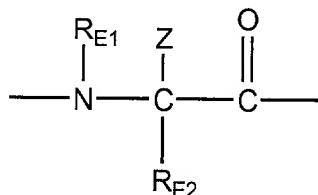
- where one of R_{B1} and R_{B2} is a substituted or unsubstituted linear, branched or cyclic alkyl, aryl, arylalkyl, heteroaryl or heteroarylalkyl group; and the other is hydrogen. X is hydrogen, methyl, trifluoromethyl or halogen, such as fluorine, chlorine, bromine or
- 5 iodine.

D is a D- amino acid residue or N-substituted glycyl residue of the formula



- where one of R_{D1} and R_{D2} is a substituted or unsubstituted aryl, heteroaryl, arylmethyl, heteroarylmethyl, benzo-fused aryl, benzo-fused heteroaryl, benzo-fused arylmethyl;
- 10 benzo-fused heteroarylmethyl, cycloalkyl or bicycloalkyl; and the other is hydrogen. Y is hydrogen, methyl, trifluoromethyl or halogen, such as fluorine, chlorine, bromine or iodine.

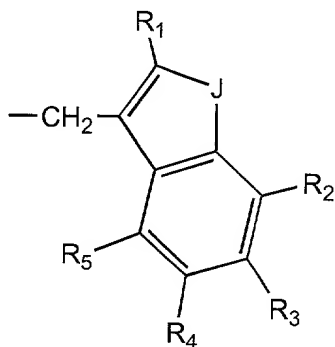
E is a D-amino acid residue or N-substituted glycyl residue of the formula



- 15 where one of R_{E1} and R_{E2} is a substituted or unsubstituted, linear, branched or cyclic alkyl, aryl or arylalkyl group; and the other is hydrogen. Z is hydrogen, methyl, trifluoromethyl or halogen, such as fluorine, chlorine, bromine or iodine.

K, L, M and N are each, independently, composed of from 1 to about 6 (which can be the same or different), D-amino acid residues, L-amino acid residues, N-substituted glycyl residues or a combination thereof. Natural or nonnatural amino acid residues can be used. One or more of the amino acid residues or N-substituted glycyl residues can, optionally, be substituted at the α -carbon by a methyl or trifluoromethyl group, or a halogen, such as a fluorine, chlorine, bromine or iodine atom.

In a preferred embodiment, one of R_{A1} and R_{A2} and one of R_{D1} and R_{D2} are, independently, a phenyl, substituted phenyl, naphthyl, substituted naphthyl, naphthylmethyl, substituted naphthylmethyl, benzyl or substituted benzyl group, or a group of the formula



where J is O, S or NR, where R is H or linear, branched or cyclic C_1 - C_6 -alkyl, preferably methyl. R_1 , R_2 , R_3 , R_4 and R_5 are independently selected from the group consisting of hydrogen, halogen and alkyl, preferably, linear, branched or cyclic C_1 - C_4 -alkyl, such as methyl. Suitable phenyl, naphthyl, naphthylmethyl and benzyl substituents include alkyl, preferably linear, branched or cyclic C_1 - C_4 -alkyl, such as methyl; and halogen, such as fluorine, chlorine, bromine or iodine. More preferably, R_{A1} and R_{D1} are both hydrogen, and R_{A2} and R_{D2} are each, independently, one of the foregoing groups.

Preferably, one of R_{B1} and R_{B2} is hydrogen, substituted or unsubstituted linear, branched or cyclic C_1 - C_4 -alkyl, phenyl, benzyl, naphthyl or naphthylmethyl. Suitable substituents include linear, branched or cyclic C_1 - C_4 -alkyl groups and halogens, such as

fluorine, chlorine, bromine or iodine. More preferably, R_{B1} is hydrogen and R_{B2} is one of the foregoing groups.

Preferably, one of R_{E1} and R_{E2} is a substituted or unsubstituted, linear, branched or cyclic C_1 - C_6 -alkyl group or a substituted or unsubstituted phenyl or naphthyl group.

- 5 Suitable substituents include linear, branched or cyclic C_1 - C_4 -alkyl groups, such as methyl, and halogens, such as fluorine, chlorine, bromine and iodine. More preferably, R_{E1} is hydrogen and R_{E2} is one of the foregoing groups.

In a preferred subset of the compounds of formula I, A and D are each a D-tryptophan residue and E is a D-leucine residue.

- 10 Preferably, K is a D-amino acid residue or an N-substituted glycyl residue comprising an amino-, carboxyl- or sulfhydryl substituted side chain, such as a cysteine, glutamic acid, aspartic acid or lysine residue, and L is a polypeptide comprising 2 or 3 D-amino acid residues, L-amino acid residues (the D- or L-amino acid residues can be the same or different) or N-substituted glycine residues. For example, in one
- 15 embodiment, L comprises 2 or 3 residues selected from among D-glycine, D-alanine or D- α - C_1 - C_4 -alkylglycine.

- Preferably, M is a polypeptide group comprising from 2 to about 8 D-amino acid residues, of which at least one comprises an amino-, carboxy- or sulfhydryl substituted side chain, such as a cysteine, glutamic acid, aspartic acid or lysine residue. N is,
- 20 preferably, a polypeptide group comprising from 1 to about 6 amino acid residues, of which at least one is a lysine residue.

- The identity of divalent linking group F is not critical, as long as it is of a suitable length to position residues A to E to interact with the N-helix coiled coil cavity (J.R. Morphy, *Curr. Op. Drug Discov. Develop.*, 1:59-65 (1998)). For example, F
- 25 preferably has a length from about 2 to about 40 atoms. In one embodiment, F is a direct bond or a polypeptide linking group of the formula $-P_n-$, wherein n is 1 to about 12 and each P is independently an L- or D- amino acid or N-substituted glycyl residue, a glycyl residue or an N-substituted glycyl derivative.

In another embodiment, F is a substituted or unsubstituted C₄-C₄₀-alkylene group, such as a polymethylene group of the formula $-(CH_2)_m-$, wherein m is from about 4 to about 40; an alkylene group which is interrupted at one or more points by a heteroatom, such as a nitrogen, oxygen or sulfur atom. For example, F can be a group
 5 $(CH_2CH_2O)_q-$, wherein q is from 1 to about 20. F can also be an alkylene group which is interrupted at one or more points by a phenylene or heteroarylene group, or a polysaccharide group, for example, a glycoside or poly(glycoside) group comprising one or more glycoside groups, for example, from 1 to about 10 glycoside groups. Suitable glycosides include glucoside, lactoside, mannoside, galactoside, fucoside, fructoside,
 10 guloside, alloside, altroside, taloside, idoside and others, such as pyranosides and furanosides, which are known in the art.

In compounds of Formula I having a C-terminal amino acid residue, the C-terminal residue can be, for example, in the form of an amide, an N-substituted amide or a carboxylic acid protecting group, as is known in the art. The nitrogen atom of an N-
 15 terminal residue can be acylated, for example, acetylated, or substituted with an amino protecting group, as is known in the art.

The term "D-amino acid residue", as used herein, refers to an α -amino acid residue having the same absolute configuration as D-glyceraldehyde. When the amino acid residue includes a first non-hydrogen α substituent and a second α substituent
 20 selected from methyl and halogen, the absolute configuration is the same as that of D-glyceraldehyde with the second α substituent taking the place of the hydrogen atom at the glyceraldehyde α -carbon.

The peptides, portions of the peptides, variations/derivatives of the peptides or portions of the variations/derivatives described herein can be used as inhibitors of HIV
 25 entry into cells. The peptides represented in Figure 3 or a portion of a peptide sufficient to fit into the hydrophobic pocket at the C-terminal end of the coiled-coil and prevent interaction of the C-peptide region with the N-peptide region of gp41 are useful to inhibit HIV infection. A portion of any of the peptides represented or of a derivative thereof can be from 2 to 20 (any number of residues from 2 to 20) amino acid residues

in size. D-peptides which comprise the consensus sequence tryptophan-tryptophan-leucine or the sequence tryptophan-tryptophan-leucine-glutamate, described herein, and additional residues, can be used; the other residues present in such D-peptides and the size of the D-peptides can be selected with reference to peptides described herein or can

5 be designed independent of those peptides, provided that these three or four residues are positioned in such a manner that the peptide can fit into the hydrophobic pocket and act as an inhibitor. Additional amino acid residues can also be present at the N-terminus, the C-terminus or both of the D-peptides described herein, thus producing a larger peptide. Alternatively, there can be other amino acid residues selected, for example, to

10 enhance binding affinity. Alternatively, a peptide which comprises the conserved amino acid residues of the D-peptides of Figure 3 can be used. For example, such a peptide can be 16 amino acid residues in size and include the conserved amino acid residues, which can be at the same positions as those at which they occur in the peptides shown in Figure 3. The intervening amino acid residues can be different from the amino acid

15 residues at these positions in any of the peptides shown in Figure 3 (e.g., can be isoleucine or asparagine or other amino acid residue which does not appear in the peptides represented in Figure 3) or can be substituted for or replaced by an amino acid residue represented at a specific position in another peptide shown in Figure 3 (e.g., the aspartic acid residue in D10pep1 can be replaced by a serine residue). Amino acid

20 residues other than the D-versions of the 20 L-amino acids found in natural proteins can be used. Such changes can be made, for example, to enhance bioavailability, binding affinity or other characteristic of the peptide. A D-peptide can comprise the conserved amino acid residues present in the peptides shown in Figure 3, but they can be separated by fewer (or more) amino acid residues than the number of intervening amino acid

25 residues shown in Figure 3. For example, fewer than five amino acid residues (e.g., Tarrago-Litvak, L. *et al.*, *FASEB, J.*, 8:497 (1994); Tucker, T.J. *et al.*, *Methods Enzymol.*, 275:440 (1996), Tarrago-Litvak, L. *et al.*, *FASEB, J.*, 8:497 (1994); Tucker, T.J. *et al.*, *Methods Enzymol.*, 275:440 (1996)), can be present between the first cysteine and the glutamic acid in the consensus sequence shown in Figure 3. Alternatively, these

two residues can be separated by more than five amino acid residues. Internal modifications can also be made (e.g., to enhance binding or increase solubility of a peptide). For example, the first tryptophan of D10pep5 can be replaced by an arginine to increase solubility. A D-peptide can have additional moieties or amino acids at its N-terminus. For example, a moiety which blocks the N terminus or gets rid of the charge otherwise present at the N-terminus can be added. The moiety can be, for example, a blocking moiety, such as an acetyl group linked directly to the glycine (G), or an acetyl group linked to one or more additional amino acid residues linked to the N-terminal of G, such as an acetyl group linked to one or more lysine residues, which, in turn, are linked to the N terminal G. In one embodiment, two lysine residues are linked to the N-terminal G (KKGAC), for example to increase the solubility of the peptide; a blocking moiety, such as an acetyl group, can be linked to the terminal lysine (acetyl group KKGAC). In another embodiment, four lysine residues are linked to the N-terminal G. In addition, a D-peptide can have additional and/or altered moieties or amino acids at its C-terminus. For example, one or both of the alanine residues at the C-terminus can be altered and/or one or more residues can be added at the C-terminus, for example to enhance binding. Alternatively, functional (chemical) groups other than amino acid residues can be included to produce an inhibitor of the present invention. For example, these additional chemical groups can be present at the N-terminus, the C-terminus, both termini or internally. In addition, two or more D-peptides can be linked via an appropriate linker (e.g., a linker of amino acid residues or other chemical moieties) to increase the effectiveness of inhibition. Alternatively, one or more D-peptides can be linked via an appropriate linker to a molecule (drug) that binds to HIV gp120, CD4, CCR5, CXCR4, or a non-pocket region of HIV gp41 to increase the effectiveness of inhibition.

The D-peptides (or L-peptides or peptides with both D- and L-amino acids) can be produced using known methods, such as chemical methods or recombinant technology. The polypeptide backbone can be altered (e.g., N-methylation) or replaced with alternative scaffolds (e.g., peptoids) at one or more positions of the peptides.

Additional components can be included in the peptides, such as, for example, linkers (chemical, amino acid) which are positioned between amino acids or amino acid portions of the peptide (e.g., to provide greater flexibility or to provide greater rigidity). As described herein, the D-peptides of the present invention are flanked by GA at the N-terminus and AA at the C-terminus, due to the design of the library used in identifying the D-peptides. Some or all of these four amino acid residues may be altered, replaced or deleted in order to produce D-peptides with, for example, altered absorption, distribution, metabolism and/or excretion. In one embodiment, the C-terminus is modified by the addition of a glycine residue immediately before the C-terminal amide.

10 In another embodiment, the most C-terminal A is altered/modified or replaced by a different amino acid residue or deleted.

D-peptides, which are of the opposite handedness from the handedness of naturally-occurring peptides, do not serve as efficient substrates for enzymes, such as proteases, and, therefore, are not as readily degraded as L-peptides. In addition, there is no effective immune response which targets D-peptides and therefore, they do not elicit an immune response comparable to that elicited by L amino acid peptides.

The present invention is illustrated by the following examples, which are not intended to be limiting in any way.

Example 1 Synthesis of Variants of the C34 Peptide

20 Mutant peptides were synthesized by solid-phase Fmoc peptide chemistry and have an acetylated amino terminus and an amidated carboxy terminus. After cleavage from the resin, peptides were desalted with a Sephadex G-25 column (Pharmacia), and then purified by reverse-phase high-performance liquid chromatography (Waters, Inc.) on a Vydac C18 preparative column using a linear water-acetonitrile gradient and 0.1% trifluoroacetic acid. Peptide identities were verified by MALDI mass spectrometry (Voyager Elite, PerSeptive Biosystems). Peptide concentrations were measured by tryptophan and tyrosine absorbance in 6 M GuHCl [H. Edelhoch, *Biochemistry*, 6:1948 (1967)].

Example 2 Quantitation of Helical Content and Thermal Stability of Mutant N36/C34 Complexes

CD measurements were performed in phosphate-buffered saline (50 mM sodium phosphate, 150 mM NaCl, pH 7.0) with an Aviv Model 62DS spectrometer as

- 5 previously described (M. Lu, S. C. Blacklow, P. S. Kim, *Nat. Struct. Biol.*, 2:1075 (1995)). The apparent melting temperature of each complex was estimated from the maximum of the first derivative of $[\theta]_{222}$ with respect to temperature.

The mean residue ellipticities ($[\theta]_{222}$, $10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$) at 0°C were as follows:

- wildtype, -31.7; Met⁶²⁹→Ala; -32.0; Arg⁶³³→Ala, -30.7; Ile⁶³⁵→Ala, -25.9; Trp⁶²⁸→Ala, -
 10 27.0; Trp⁶³¹→Ala, -24.9. In the case of the Trp⁶²⁸→Ala and Trp⁶³¹→Ala mutations, the decrease in $[\theta]_{222}$ is likely to overestimate the actual reduction in helical content. The removal of tryptophan residues from model helices has been reported to significantly reduce the absolute value of $[\theta]_{222}$ even when there is little change in helical content (A. Chakrabarty, T. Kortemme, S. Padmanabhan, R. L. Baldwin, *Biochemistry*, 32:5560
 15 (1993)).

Example 3 Identification of Peptides Which Bind to a Pocket on the Surface of the N-helix Coiled-Coil of HIV-1 gp41.

Methods are available to identify D-peptides which bind to a cavity on the surface of the N-helix coiled-coil of HIV envelope glycoprotein gp41. As described in
 20 detail below, D-peptides which bind to a cavity on the surface of the N-helix coiled-coil of HIV-1 envelope glycoprotein gp41 were identified by mirror-image phage display.

- This method involves the identification of ligands composed of D-amino acids by screening a phage display library. D-amino acid containing ligands have a chiral specificity for substrates and inhibitors that is the opposite of that of the naturally
 25 occurring L-amino ligands. The phage display library has been used to identify D-

amino acid peptide ligands which bind a target or desired L-amino acid peptide (Schumacher *et al. Science*, 271:1854-1857 (1996)).

D-peptides that bind to the hydrophobic pocket of gp41 were identified using a target that is an enantiomer of IQN17, a hybrid molecule containing 29 residues of GCN4-pI_QI on the N-terminal end and 17 residues of gp41 on the C-terminus. The phage library used for selection is described in U.S. Patent 5,780,221 and Schumacher *et al. Science*, 271:1854-1857 (1996). The complexity of the library is greater than 10⁸ different sequences. The sequences are flanked on either end by either a cysteine or a serine, with ten random residues in the middle. These sequences are located in the pIII gene of the phage, a coat protein that is expressed as approximately five copies on the outer surface of the phage.

The following experimental procedures were used in the examples described herein.

Phage Display

Neutravidin (Pierce, 10 µg in 100 µL of 100 mM NaHCO₃) was added to individual wells of a 96-well high-binding styrene plate (Costar) and incubated overnight on a rocking platform at 4°C. The neutravidin was removed and the wells were washed four times with a TBS/Tween solution. Biotinylated D-IQN17 (100 µL of a 10 µL peptide solution in 100mM NaHCO₃) was added to the wells and incubated for one hour at 25°C. The biotinylated target was removed and a blocking solution (30 mg/ml nonfat dried milk in 100 mM NaHCO₃) was added to the wells and incubated for two hours, with rocking, at 4°C. The blocking solution was removed and the wells were coated again with the biotinylated target as above. The target was removed and the unliganded neutravidin was blocked by the addition of the blocking solution with 5 mM biotin. After removing the biotin, the wells were washed six times with the TBS/Tween solution. The phage stock was then added to the wells (50 µL of phage stock plus 50 µL of phage-binding buffer: TBS, 0.1% Tween-20, 1 mg/ml milk, 0.05% sodium azide). The incubation time of the phage stock in the wells decreased in increasing rounds of

selection. After incubation, the phage solution was removed and the wells were washed twelve times with TBS/Tween to remove the unbound phage. Odd numbered washes were performed quickly, with no incubation time; even numbered washes were incubated for increasing amounts of time each round of phage selection. The phage were eluted by the addition of two micrograms of trypsin in 100 μ L of phage-binding buffer and 2.5 mM CaCl_2 with an hour incubation at 37°C. To determine recovery, a dilution of the eluted phage was used to infect K 91 kan cells. After a one hour incubation, 100 μ L of cells were removed and 1:10, 1:100, and 1:100 dilutions in LB were plated on LB/tetracycline plates. Phage recovery was determined as a ratio of transducing units recovered (the titer of the eluted phage) to the input number of transducing units (the titer of the phage stock used that round). Transducing units were determined by counting the number of tetracycline-resistant colonies on the LB/tetracycline plates. Non-specific phage recovery generally has a ratio in the order of magnitude of 10^{-8} to 10^{-9} , whereas specifically amplified phage have a ratio 10^{-7} or greater. Individual clones were amplified and sequenced. They were assayed in the binding assay to determine binding specificity.

D10pep7 was identified after five rounds of phage selection. D10pep1, D10pep3, D10pep4, D10pep5, and D10pep6 were identified after seven rounds of phage selection. The phage selection was performed again, with shorter incubation times and longer washes, and D10pep10 and D10pep12 were identified after three rounds of selection. (A ninth D-peptide was identified but was not further investigated once it was shown to be toxic to cells.)

To test the specificity of binding of identified phage clones to the pocket of D-IQN17, the phage clones were added to wells of 96-well plates coated as above with D-IQN17, D-GCN4-pIQI (with the three mutations), D-IQN17(G39W = glycine36 substituted with tryptophan), or wells with no target. The phage were incubated on the plates and washed for the same lengths of time as in the round from which they were identified. Eluted phage were used to infect K91 kan cells and the recovered

transducing units were determined as above. These sequences bound specifically to the wells with D-IQN17.

Peptide Purification

IQN17 and the D10 peptides were synthesized by Fmoc peptide chemistry.

- 5 They have an acetylated N-terminus and a C-terminal amide. IQN17 contains 29 residues derived from GCN4-pI_QI on the N-terminus and 17 residues from the C-terminus of N36 on the C-terminus. There is one residue overlap between GCN4-pI_QI and the N36 region, making the peptide 45 residues long. To improve solubility, three amino-acid substitutions were made in the GCN4-pI_QI region of IQN17, as compared to
- 10 the original GCN4-pI_QI sequence (Eckert, D.M. *et al.*, *J. Mol. Biol.*, 284:859-865 1998). These substitutions are L13E, Y17K, and H18K. Thus, the sequence of IQN7 is: ac-RMKQIEDKIEEIESKQKKIENEIARIKKLLQLTVWGIKQLQARIL-am
- (ac- represents an N-terminal acetyl group and -am represents a C-terminal amide), with the HIV portion underlined. For mirror-image phage display, IQN17 was synthesized
- 15 using D-amino acids (for amino acid residues that contain a second chiral center, such as Ile and Thr, the exact mirror image of the naturally occurring amino acid residue is used to create the D-version of the target). In addition, the N-terminus of the peptide was biotinylated using NHS-LC-biotin II (Pierce, catalog #21336). Between the biotin and the IQN17 sequence was a three amino acid linker of GKG, with the lysine in the
- 20 naturally-occurring L-form. This lysine was inserted as a trypsin recognition site.

The sequences of the D-peptides are as follows (with all amino acids in the D-enantiomer, using the exact mirror image of naturally occurring amino acid residues for Ile and Thr, which contain a second chiral center):

- D10pep1: Ac-GACEARHREWLCAA-CONH₂ (SEQ ID NO: 34);
- 25 D10pep3: Ac-KKGACGLGQEEFWLCAA-CONH₂ (SEQ ID NO: 15);
- D10pep4: Ac-GACDLKAKEFWLCAA-CONH₂ (SEQ ID NO: 35);
- D10pep5: Ac-KKGACELLGWEWLCAA-CONH₂ (SEQ ID NO: 16);
- D10pep6: Ac-GACSRSPQPEWLCAA-CONH₂ (SEQ ID NO: 36);

D10pep7: Ac-GACLLRAPEWGWLCAA-CONH₂ (SEQ ID NO: 37);
 D10pep10: Ac-KKGACMRGEWEWSWLCAA-CONH₂ (SEQ ID NO: 18); and
 D10pep12: Ac-KKGACPPLNKEWAWLCAA-CONH₂ (SEQ ID NO: 19).

After cleavage from the resin, the peptides were desalted on a Sephadex G-25
 5 column (Pharmacia) and lyophilized. The lyophilized peptides were purified by
 reverse-phase high performance liquid chromatography (Waters, Inc.) on a Vydac C18
 preparative column. The D-peptides were then air-oxidized by dissolving the
 lyophilized powder in 20 mM Tris, pH 8.2, and stirring at room temperature for several
 days. The oxidized peptides were HPLC purified as before. The expected molecular
 10 weights of the peptides were verified using MALDI-TOF mass spectrometry
 (PerSeptive Biosystems). Peptide concentrations were determined using tyrosine,
 tryptophan and cysteine absorbance at 280 nm in six molar GuHCl (Edelhoch, 1967).
 Peptide stock solutions were prepared in DMSO.

The N-terminal lysines on D10pep3, D10pep5, D10pep7a, D10pep10 and
 15 D10pep12 were added to increase the water solubility of the peptides. To investigate
 the effect of the added lysines on the inhibitory activity of the peptides, D10pep1 was
 synthesized with two N-terminal lysines (denoted D10pep1a) and compared to D10pep1
 without lysines: D10pep1a was found to have an IC₅₀ for inhibition of syncytia
 formation approximately 2-fold higher than D10pep1 (i.e., without lysines). In addition,
 20 D10pep5 was synthesized with two additional N-terminal lysines (for a total of four
 lysines to generate a peptide denoted D10pep5a). The IC₅₀ for inhibition of syncytia
 formation of D10pep5a was approximately 2-fold higher than D10pep5. The addition
 of N-terminal lysine residues to the D-peptides results in only a modest decrease of
 inhibitory activity.

25 D-peptides that had additional D-Lys residues added to the N-termini, that were
 synthesized for study are indicated with the addition of "a" to the peptide name and
 include the following:

D10pep1a: Ac-KKGACEARHREWAWLCAA-CONH₂ (SEQ ID NO: 38);

- D10pep4a: Ac-KKGACDLKAKEWFWLCAA-CONH₂ (SEQ ID NO: 39);
D10pep5a: Ac-KKKKGACELLGWEWAWLCAA-CONH₂ (SEQ ID NO: 17);
D10pep6a: Ac-KKGACSRSQPEWEWLCAA-CONH₂ (SEQ ID NO: 40); and
D10pep7a: Ac-KKGACLLRAPEWGWLCAA-CONH₂ (SEQ ID NO: 41).

5 These sequences are also represented in Figure 3. The 12 amino acid “core” of each D-peptide (which, in turn comprises a 10-mer and the consensus sequences described herein) are as follows:

	CDLKAKEWFWLC	(SEQ ID NO: 3)
	CEARHREWAWLC	(SEQ ID NO: 4)
10	CELLGWEWAWLC	(SEQ ID NO: 5)
	CLLRAPEWGWLC	(SEQ ID NO: 6)
	CSRSQPEWEWLC	(SEQ ID NO: 7)
	CGLGQEEWFWLC	(SEQ ID NO: 8)
	CMRGEWEWSWLC	(SEQ ID NO: 9)
15	CPPLNKEWAWLC	(SEQ ID NO: 10)
	CVLKAKEWFWLC is an alternative sequence for peptide SEQ ID NO: 3. (SEQ ID NO: 11).	

It is readily apparent that there is a highly conserved consensus sequence in these peptides. The 12 amino acid peptide represented in Figure 3 can be represented as:
20 CXXXXXEWXWLC (SEQ ID NO: 12), where amino acid residues common to the peptides are shown and X represents an amino acid residue which is not conserved among the peptides.

Example 4 Assessment of Activity of C34 Peptides and D-Peptides

The potency of C34 peptides in inhibiting viral infection and the HIV-1 infection
25 inhibitory activity of the D-peptides were assayed using recombinant luciferase-expressing HIV-1 (Chen, B.K. *et al.*, *J. Virol.*, 68:654 (1994); Malashkevich, V.N., *et al. Proc. Natl. Acad. Sci., USA*, 95:9134 (1998)). The virus was produced by co-

transfecting an envelope-deficient HIV genome NL43LucR-E- (Chen, B.K. et al., *J. Virol.*, 68:654 (1994) and the HXB2 gp160 expression vector pCMVHXB2gp160 (see Chan, D.C. et al., *Proc. Natl. Acad. Sci.*, 95:11513 (1998)) into 293T cells. Low-speed centrifugation was used to clear the viral supernatants of cellular debris. The

5 supernatant was used to infect HOS-CD4/Fusion cells (N. Landau, NIH AIDS Reagent Program) in the presence of the D-peptides, with concentrations ranging from 0 to 500 μ M. Cells were harvested 48 hours post-infection, and luciferase activity was monitored in a Wallac AutoLumat LB953 luminometer (Gaithersburg, MD). The IC_{50} is the peptide concentration that results in a 50% decrease in activity relative to control

10 samples lacking peptide. The IC_{50} was calculated from fitting the data to a Langmuir equation $[y=k/(1+([peptide]/IC_{50}) + x)]$, where y = luciferase activity and k and x are scaling constants.

Cell/Cell Fusion Assay

Inhibition of cell/cell fusion (i.e., syncytia formation) was assayed by co-

15 culturing Chinese hamster ovary cell expressing HXB2 envelope (K. Kozarsky, et al., *J. Acquir. Immune. Defic. Syndr.*, 2:163 (1989) and the HeLa-CD4-LTR-Beta-gal cells (M. Emerman, NIH AIDS Reagent program) in the presence of varying concentration of peptide. When mixed, these cells form syncytia, or multi-nucleated cells, which express β -galactosidase. Approximately twenty hours after co-culturing the cells, the

20 monolayers were stained with 5-bromo-4-chloro-3-indolyl- β -D-galactoside to visualize the syncytia. The syncytia are visualized with a microscope and counted manually (a syncytia is scored as a fused cell containing three or more nuclei). The IC_{50} was calculated from fitting the data to a Langmuir equation $[y = k/(1 + [peptide]/IC_{50}) + x]$, where y = number of syncytia and k and x are scaling constants.

Table 1 Stability of mutant N36/C34 complexes and the inhibitory potency of C34 mutants.

Peptide	T _m (°C)	IC ₅₀ (nM) viral entry	IC ₅₀ (nM) cell fusion
Wildtype	66	2.1 ± 0.31	0.55 ± 0.03
<u>Cavity-binding</u>			
Trp ⁶²⁸ _Ala	53	10 ± 2.0	3.8 ± 0.33
Trp ⁶³¹ _Ala	37	61 ± 16	15 ± 0.82
Ile ⁶³⁵ _Ala	55	4.1 ± 0.91	0.96 ± 0.12
<u>Control residues</u>			
Met ⁶²⁹ _Ala	66	2.0 ± 0.27	0.74 ± 0.03
Arg ⁶³³ _Ala	65	2.6 ± 0.89	0.76 ± 0.07

Mutant C34 peptides (10 μM) were complexed with the N36 peptide (10 μM) in phosphate-buffered saline (pH 7.0) for circular dichroism (CD) measurements. The apparent melting temperatures (T_m) were estimated from the thermal dependence of the CD signal at 222 nm. Inhibition of viral entry was measured in a cell-culture infection assay using recombinant luciferase-expressing HIV-1. Inhibition of cell-cell fusion was measured in a syncytium assay. The means and standard errors are from triplicate trials.

Similarly, the activity of the D-peptides described was assessed using the two assays described above. Results are shown in figures 6A-6B and 8A-8B.

Example 5: Crystallization of the IQN17/D10pep1 Complex and Ligand-Free IQN17

Peptide Purification, Crystallization

5 Peptides IQN17 and D10pep1 were synthesized by Fmoc peptide chemistry, as described above.

 A 10 mg/ml stock of a mixture of IQN17 and D10pep1 was prepared in water. The final concentration of IQN17 was about 1.37 nM, and the final concentration of D10pep1 was about 1.51 mM. Initial crystallization conditions were found using

10 Crystal Kits I and II (Hampton Research), and then optimized. To grow the best diffracting crystals, one microliter of this stock was added to one microliter of the reservoir buffer (10% PEG 4000, 0.1 M NaCl pH 5.6, 20 % 2-propanol) and allowed to equilibrate against the reservoir buffer. Crystals belong to a space group P321 ($a=b=41.83\text{\AA}$; $c=84.82\text{\AA}$, $\alpha=\beta=90^\circ$, $\gamma=120^\circ$) and contain one IQN17/D10pep1 monomer

15 in the asymmetric unit. A useful osmium derivative was produced by increasing the concentration of PEG 4000 in the reservoir solution by 4%, adding $(\text{NH}_4)_2\text{OsCl}_6$ to the reservoir solution to a final concentration of 5 mM and adding five microliters of the resulting solution to the drop containing the protein crystal. Prior to data collection native and heavy-atom derivative crystals were transferred into cryosolution containing

20 20% PEG 4000, 0.1 M NaCl PH 5.6, 20% 2-Propanol and flash-frozen using X-stream cryogenic crystal cooler (Molecular Structure Corporation).

 The best diffracting crystals of ligand-free IQN17 were grown with a similar technique as above: on microliter of a 10 mg/ml solution of IQN17 in water was added

to one microliter of the reservoir buffer (1.0 M K₂Na Tartrate, 0.1 M NaHEPES pH 7.0) and allowed to equilibrate against the reservoir buffer. Before flash freezing, the crystals were transferred into buffers consisting of the reservoir solution with increasing amounts of glycerol, up to a final concentration of 23% glycerol. Crystals belong to the

5 space group C222₁ ($a=57.94\text{ \AA}$, $b=121.96\text{ \AA}$, $c=73.67\text{ \AA}$; $\alpha=\beta=\gamma=90^\circ$) and contain one IQN17 trimer in the asymmetric unit.

X-Ray Data Collection and Processing

Initial data were collected on a Rigaku RU300 rotating-anode x-ray generator mounted to an R-axis IV area detector (Molecular Structure Corporation). Diffraction

10 data for IQN17 were collected at 100 K using a Quantum-4 CCD detector and the 5.0.2 beamline at the Advanced Light Source (Berkeley, USA). Final native and multiwavelength anomalous diffraction (MAD) data for IQN17/D10pep1 were collected at the Howard Hughes Medical Institute Beamline X4A at Brookhaven National Laboratory using a Raxis-IV detector. For MAD data, four wavelengths near the

15 osmium L-III absorption edge were selected based on the fluorescence spectrum of the Os derivative crystal (Table 2). The four wavelengths were: 1.1398 \AA , 1.1403 \AA , 1.1393 \AA , 1.1197 \AA . Data sets were collected in 20° batches, allowing the same batch to be collected at each wavelength before moving to the next batch, in order to minimize the crystal decay between data sets. Reflections were integrated and scaled with the

20 programs DENZO and SCALEPACK (Otwinowski, Z., (1993) in *Data Collection and Processing*, eds. Sawyer, L., Isaacs, N. & Bailey, S. (SERC, Daresbury Laboratory, Warrington, England), pp. 55-62).

Further diffraction data processing, phase determination and map calculations were performed using the CCP4 suite of programs (*CCP4, Acta Cryst. D50:760-763*

25 (1994)). Intensities were reduced to amplitudes with the program TRUNCATE, and the data sets for the wavelengths closest to the Os L-III absorption edge (λ_1 , λ_2 , λ_3) were scaled with SCALEIT to the remote wavelength (λ_4) data set (Table 2).

Phase Determination and Crystallographic Refinement

Initially, phase determination for IQN17/D10pep1 crystals was attempted with the molecular replacement technique using the theoretical model of IQN17 build from the published GCN4-pIQI and HIV gp41 structures (Eckert, D.M., *et al.* (1998) *J. Mol. Biol.* 284:859-865; Chan, D.C., *et al.* (1997) *Cell* 89, 263-273) with sidechains truncated to a polyserine chain. The resulting molecular replacement solutions were ambiguous and the electron density map did not reveal conformation of the D10pep1 peptide. The molecular replacement phases were good enough, however, for determining the coordinates of a single Os atom in the corresponding derivative using difference and anomalous fourier maps. The heavy atom binds on the cryallographic three-fold axis (0.333, 0.667, 0.047). MAD phases were then generated with the program MLPHARE (Table 2) and extended to higher resolution with the program DM. The quality of MAD electron density map at 1.5 Å resolution was exceptional, and revealed structural details of IQN17 and D10pep1 peptide with clarity. Electron density map interpretation and model building was done with the program O (Jones, T.A. *et al.* (1991) *Acta Crystallogr.* D47, 110-119). The structure of IQN17-D10pep1 complex was refined using the program CNS (Brünger, A.T. *et al.*, *Acta Crystallogr.* D54, 905-921 (1998)). The correctness of the structure was checked with simulated annealing omit maps and with the program WHAT CHECK (Hoff, R.WW. *et al.*, *Nature* 381: 272 (1996)). All residues of IQN17 and the D10pep1 peptide (when converted into its mirror image) occupy most preferred areas of the Ramachandran plot. The conformations of the majority of the residues are well defined except for the two most N-terminal residues of IQN17 and the side chains of Arg-6 and Arg-8 of the D10pep1 peptide.

The structure of ligand-free IQN17 was solved by molecular replacement using the program AMORE (Navaza, J. (1994) *Acta Crystallogr.* A50, 157-163) and the IQN17 part of the refined IQN17/D10pep1 structure as a test model. Three-fold noncrystallographic averaging, solvent flattening and histogram matching with the program DM was used for phase improvement. Electron density map interpretation and

model building was done with the program O (Jones *et al.*, *Acta Crystallogr. D* 54, 905-921 (1991)). The structure of the IQN17/D10pep1 complex was refined using the program CNS (Brunger, A.T. *et al.*, *Acta Crystallogr. D* 54, 905-921 (1998)).

The crystal structure can be used to design more effective and/or new D-peptides, peptidomimetics or other small molecules that inhibit HIV infectivity.

Example 6 Nuclear magnetic resonance (NMR) methods for identifying compounds which bind to the N-helix hydrophobic pocket of gp41

A. Assaying specific binding between the IQN17 hydrophobic pocket and D-peptides

NMR experiments were used to assay the binding of each D-peptide to IQN17. The single tryptophan residue of IQN17 (denoted Trp-571) provides an excellent probe of specific binding to the hydrophobic pocket of gp41. In deuterium oxide (deuterated water) buffers, the simple homonuclear one-dimensional ^1H NMR spectrum of IQN17 (Figure 9A, middle) shows five signals from the Trp-571 indole, extremely well-resolved from all other signals in the molecule. To test a compound for binding to the gp41 pocket, two one-dimensional ^1H NMR measurements were made on samples in deuterated buffers. First, a reference (control) spectrum of IQN17 was taken, identifying the Trp571 chemical shifts in the unbound form. A second spectrum was acquired on a sample containing both IQN17 and the compound in question. An optional third spectrum of the D-peptide (or other small molecule, or mix of molecules) was also taken. ^1H NMR experiments were performed on a Bruker AMX 500 spectrometer. Data was processed in Felix 98.0 (MSI) on Silicon Graphics computers, and all spectra were referenced to DSS. All experiments were performed at 25° C in 100 mM NaCl, 50 mM sodium phosphate (pH 7.5). All buffers used were >99.7% D_2O , to remove overlapping resonances from exchangeable backbone and side chain protons. Solute concentrations ranged from 0.3-1.0 mM for individual peptides, 0.8-1.0 mM for 1:1 complexes of IQN17 with each D-peptide.

Simple binding of two or more components is expected to result both in broader peaks (due to the increased size of the complex) and in changes in chemical shifts (due to the different chemical environments experienced by nuclei in free and bound forms). Specific binding to the hydrophobic pocket is indicated by a change in the Trp-571
 5 chemical shifts, as well as by a broadening of peaks. Binding can also be indicated by similar changes in the chemical shifts and peak widths of the molecule (peptides and small organic molecules, for example) assayed. Figure 9A shows an example of these effects: the NMR spectrum of the IQN17/D10pep1a complex displays broader peaks and dramatically different chemical shifts than the spectra for either of the two separate
 10 components. All IQN17/D-peptide complexes studied gave similar results, though varying in the degree of chemical shift dispersion (Figure 9B). Thus, binding was indicated in all cases.

The x-ray crystallographic finding that the two conserved Trp residues, and the conserved Leu residue, in D10pep1 are directly involved in the binding of the IQN17
 15 pocket, strongly suggests that these conserved residues participate in a similar manner when the other D-peptides bind the pocket. These conserved tryptophan residues, and Trp-571 of IQN17, provide an opportunity to study the binding interfaces in greater detail. In the IQN17/D10pep1 crystal structure, the Trp-571 sidechain of IQN17 is in close contact with Trp-10 of D10pep1, with several protons of Trp-571 ($H_{\zeta 2}$, $H_{\eta 2}$, $H_{\zeta 3}$,
 20 $H_{\epsilon 3}$; the four scalar-coupled protons of the aromatic ring) above the plane of the Trp-10 indole group. In this position, aromatic ring current interactions (F.A. Bovey, *Nuclear Magnetic Resonance Spectroscopy* (1988)) are expected to alter the chemical shifts of some of those protons, moving peaks upfield in the manner seen (Figure 9A, bottom). Use of the structure-based chemical shift prediction program SHIFTS (version 3.0b2, K.
 25 Osapay, D. Sitkoff, D. Case) also predicted that only protons from Trp-571 will experience a large upfield shift, especially the $H_{\zeta 3}$ proton. If the other D-peptides bind to the IQN17 pocket in the same fashion as D10pep1, a similar juxtaposition of Trp-571 and Trp-10 should occur, resulting in upfield-shifted peaks. All of the D-peptide/IQN17 complexes studied displayed such peaks, though varying in the extent of the shift

(Figure 9B). The D10pep1 complex showed the most extreme upfield shifts, and the D10pep7a complex the least. The magnitude of these changes is very large, ranging from roughly 0.5 to 2 ppm for the most upfield-shifted proton ($H_{\zeta 3}$, in all cases where it could be assigned). In comparison, chemical shift differences often used to detect

5 binding in SAR by NMR experiments (Shuker, S.B., Hajduk, P.J., Meadows, R.P., Fesik, S.W., *Science* 274:1531-1534 (1996)) are frequently in the range of 0.05 to 0.2 ppm.) Though a broad range of upfield chemical shifts was observed, ring-current effects can be highly sensitive to distance and orientation, so that small structural differences may give rise to substantial variations in chemical shift. (All of the upfield

10 shifts observed are consistent with the approximate orientation of Trp side chains expected from the x-ray crystal structure.) Also, the upfield-shifted peaks are somewhat broadened compared to others in these NMR spectra (most likely due to some type of exchange process) an effect particularly pronounced for the complexes with D10pep5a and with D10pep7a.

15 To confirm that the strongly upfield-shifted peaks all correspond to a single sidechain (almost certainly Trp-571), two-dimensional NMR (TOCSY) experiments were performed on each of the IQN17/D-peptide complexes. As expected, the TOCSY experiments indicate that in each complex, the strongly upfield-shifted resonances all belong to the same aromatic side chain, identified as a group of four scalar-coupled

20 protons. One example TOCSY spectrum is shown in Figure 9C. For several of the complexes studied, NOESY experiments also indicate contact between this sidechain and other (unassigned) aromatic groups, as expected from the IQN17/D10pep1 structure. Not all of the potential NOE crosspeaks could be resolved, due to intense spectral overlap in the 6.8-7.6 ppm region. 2D NOESY and TOCSY experiments as

25 described in J. Cavanaugh, W.J. Fairbrother, A.G. Palmer, N.J. Skelton, *Protein NMR Spectroscopy: Principles and Practice* (1996) were performed on samples of IQN17 and of each complex, with mixing times ranging between 30-90 ms (NOESY) and 30-70 ms (TOCSY). Spectral widths of 11,111 Hz and 5555 Hz were used in the acquisition (t_2) and indirect (t_1) dimensions, respectively. TOCSY experiments employed the DIPSI-

2rc mixing sequence (J. Cavanaugh, M. Rance, *J. Magn. Reson. Ser. A*, 105:328 (1993)).

We conclude that all D-peptides assayed clearly bind the hydrophobic pocket of IQN17. Additionally, in the majority of these IQN17 complexes (i.e., D10pep1, D10pep3, D10pep4, D10pep6, D10pep10, and D10pep12) the D-peptides contact the pocket with very similar binding interfaces, bringing Trp-571 in close contact with the aromatic ring of Trp-10. In the cases of complexes with D10pep5a and D10pep7a this conclusion also seems very likely, although the more limited chemical shift dispersion and broader peaks raise a remote possibility of some other mode of binding.

The binding assay employed here can also be employed to assay binding of other molecules to the hydrophobic pocket of gp41 (e.g., such as found in IQN17). The assay is especially easy to interpret in a case where an aromatic group binds the pocket, as with the set of D-peptides described above. However, any pocket-binding molecules should also perturb the chemical shifts of Trp-571, an easily noticeable effect. In addition, new NMR signals generated by the small molecules themselves upon binding, are also indicative of binding.

The use of one-dimensional homonuclear ^1H NMR provides significant advantages over multidimensional heteronuclear NMR to determine specific binding: (1) Sensitivity is higher, allowing samples to be assayed more quickly; alternately the higher sensitivity makes possible the use of lower concentrations of IQN17 and of putative binding agents, allowing screening for higher-affinity compounds, and more of them simultaneously. (2) Non-isotopically labeled proteins are simpler to produce, and more cost-effective. However, two-dimensional NMR experiments, either homonuclear or heteronuclear (with ^{15}N and/or ^{13}C isotopic labeling) could also be employed.

B. Screening chemical libraries

The binding assay described in (A) above can be used to screen large numbers of compounds present in a chemical library. Simple one-dimensional homonuclear ^1H NMR experiments are sufficient to assess binding, with no requirement for isotopic labeling. Two-dimensional NMR experiments, either homonuclear or heteronuclear

(with ^{15}N and/or ^{13}C isotopic labeling) could also be employed. Single compounds can be screened one at a time in this process. However, multiple compounds can also be combined in the same assay with IQN17 (or any representation of the gp41 N-helix coiled coil) and screened simultaneously. Binding to the pocket by any component of the mixture is indicated by a change in the Trp-571 chemical shifts. NMR signals from a large number of compounds together have the potential to obscure signals from Trp-571; these signals from unbound molecules can be eliminated using pulsed field gradient techniques well known in the art. With use of these techniques and a commercially available NMR tube sample changer, the automated screening of large numbers of compounds is straightforward.

C. Evaluating the products of multiple combinatorial syntheses

The screening process described in (B) above can also be extended to take advantage of combinatorial organic synthetic methods. Such methods are currently being used to generate whole families of compounds, with each family containing a diverse number of chemically related compounds. By the simple assay described above, the products of an entire combinatorial synthesis can be screened simultaneously. If no binding is indicated, then there is no need to invest further attention in any member of that family of compounds. If binding is indicated, then a particular family of promising compounds can be targeted for more detailed investigation. Simple one-dimensional homonuclear ^1H NMR experiments are sufficient to assess binding, with no requirement for isotopic labeling. Two-dimensional NMR experiments, either homonuclear or heteronuclear (with ^{15}N and/or ^{13}C isotopic labeling) could also be employed.

TABLE 2

Data collection								
Crystal	λ (Å)	Completeness (%)	R_{sym}^1 (%)	Resolution (Å)				
IQN17	1.0000	89.5	3.7	2.1				
IQN17/D10	1.1197	93.8	4.8	1.5				
Os λ_1	1.1403	98.6	6.3	2.0				
Os λ_2	1.1399	96.8	9.7	2.0				
Os λ_3	1.1393	96.9	7.9	2.0				
Os λ_4	1.1197	97.0	8.4	2.0				

MAD phasing statistics (22.0-2.0 Å)								
Derivative	R_{iso}^2 (%)	R_{cullis}^3		R_{cullis}^3 Anom.	Ph. Power ⁴		Occ. ⁵	Anom. Occ. ⁵
		Acentric	Centric		Acentric	Centric		
Os λ_1 vs. λ_4	7.3	0.75	0.61	0.47	1.41	1.21	-0.039	0.337
Os λ_2 vs. λ_4	5.2	0.83	0.71	0.44	1.04	1.15	-0.027	0.533
Os λ_3 vs. λ_4	3.3	0.97	0.97	0.49	0.35	0.28	-0.005	0.295

Overall figure of merit (before solvent flattening): 0.68

Refinement statistics									
Crystal	Non-hydrogen protein atoms	Water s	Ions	Resolution (Å)	Reflections total	R_{cryst}^6 R_{free}^6		R.m.s. deviations	
								bonds (Å)	angles (°)
IQN17/D10	516	150	1	10.0 - 1.5	13549	0.214	0.245	0.012	1.498
IQN17	1143	160	1	5.0 - 2.5	7541	0.282	0.352	0.009	1.252

¹ $R_{\text{sym}} = \sum \sum |I_j - \langle I \rangle| / \sum \sum \langle I \rangle$, where I_j is the recorded intensity of the reflection j and $\langle I \rangle$ is the mean recorded intensity over multiple recordings.

² $R_{\text{iso}} = \sum |F(\lambda_i) \pm F(\lambda_4)| - |F(\lambda_i)| / \sum |F(\lambda_4)|$, where $F(\lambda_i)$ is the structure factor at wavelength λ_i and $F(\lambda_4)$ is the structure factor at the reference wavelength λ_4 .

³ $R_{\text{cullis}} = \sum |F(\lambda_i) \pm F(\lambda_4)| - |F_{\text{h}(\lambda_i), \text{c}}| / \sum |F(\lambda_i) \pm F(\lambda_4)|$, where $F_{\text{h}(\lambda_i), \text{c}}$ is the calculated heavy atom structure factor.

⁴Phase power = $\langle F_{\text{h}(\lambda_i)} \rangle / E$, where $\langle F_{\text{h}(\lambda_i)} \rangle$ is the root-mean-square heavy atom structure factor and E is the residual lack of closure error.

⁵Occupancies are values output from MLPHARE.

⁶ $R_{\text{cryst, free}} = \sum |F_{\text{obs}}| - |F_{\text{calc}}| / |F_{\text{obs}}|$, where the crystallographic and free R factors are calculated using the working and test sets, respectively. Test set contained 10% of reflections.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.

0399.1192-008